Optimum isarithm interpolation in digital modelling: an examination of the performances of some computer contouring and interpolation methods and techniques applicable within digital terrain modelling

Grassie, Donald Norman Duncan

OPTIMUM ISARITHM INTERPOLATION IN DIGITAL MODELLING:

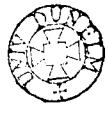
An examination of the performances of some computer contouring and interpolation methods and techniques applicable within digital terrain modelling.

Donald Norman Duncan Grassie

A thesis submitted for the degree of
Doctor of Philosophy of
the University of Durham

Department of Geography

March 1984



ABSTRACT

Surface interpolation and, more particularly, isarithm interpolation are common procedures in the Earth Sciences and increasingly common ones in most other disciplines. Data for such interpolation are increasingly being stored in a computer-accessible form as a prerequisite to a computer-derived solution. This has resulted in the development of a large assortment of computer software to perform the interpolation. The scientist is thus faced with a predicament of which is the optimum method of isarithm interpolation for his data. This thesis describes an empirical evaluation of the performances of a comprehensive selection of software in relation to data characteristics. Much of the software is available in the academic environment or was developed by the author.

The evaluation utilises detailed, photogrammetrically-derived digital terrain data which are related to some currently used data classification systems. Additionally, several techniques, mostly derived from geomorphometry, are used to examine the data in more depth, in an attempt to establish reliable descriptors of their data characteristics. These descriptors are designed to be used in a multi-disciplinary environment.

The interpolation process is essentially classified into three stages for evaluation. Random-to-grid algorithms are evaluated by considering the interpolated grid in relation to an observed grid. Grid-to-isarithm and random-to-isarithm interpolation are evaluated by considering the interpolated isarithms with observed isarithms. In both cases, geometric errors and the morphological trueness of the interpolation product are considered using graphical and numerical parameters. These statements of accuracy are correlated with the data characteristics to establish the main factors influencing isarithm interpolation.

No part of the material contained in this thesis has previously been submitted for a degree in this or any other university.

Signed: 

Date: 23rd March, 1984.

For

Oonagh and Duncan,

Mum, Dad, Catherine and Vivien.

Thank you.

ACKNOWLEDGEMENTS

This thesis reports on work carried out in the Department of Geography at Durham University, in the Department of Surveying at Newcastle University and finally in Exploration Data Processing at Britoil plc. I am grateful for the provision of the facilities in Durham enabling the initial period of research to be undertaken, and to the Social Science Research Council for a Research Studentship over that brief period. I am indebted to Mr. P.J. Carmody and my former colleagues at Newcastle for providing additional resources and encouraging the continuation of the research. Finally, my thanks to Mr. D. Stone for providing facilities at Britoil.

The further I have progressed with this thesis, the more I have realised the importance of my family to its successful completion. My mother, father and sisters, who gave me the initial quest for knowledge, have given me continual support and encouragement throughout my protracted education. My devoted wife, Oonagh, who has sacrificed much over the last few years, has augmented this support and has also helped with the preparation of this manuscript. I thank them and my son for providing the continual impetus.

I am indebted to the computing staffs at various sites where I have conducted this research. Mr. B. Lander and the operators in the Durham University Computer Unit gave me considerable assistance particularly with the plotting of diagrams. The operations staff at Newcastle and Britoil have also given me much help.

At other sites, I have been overwhelmed by the support I have received. My thanks to Professors Thompson and Petrie and the 'Topo' staff in the Department of Geography at Glasgow University who provided the resources when I collected my data and have encouraged me throughout my career. Mr. P. Smith and staff at the Rutherford Laboratories and Dr. Erskine and staff at the St. Andrews University Computing Laboratory provided essential resources for the access of the SACM and SURFACE II GRAPHICS packages.

Finally, my thanks to all my colleagues and friends who have encouraged and assisted me throughout this research. To all in 'Skylab', the 'Old Brewery Building', '220/226', and especially Andrew Haggart and Dave Rogers who always could provide the 'reasons to be cheerful' - 'this is what we find'.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

NOTATION

Throughout this thesis, algebraic notation has been kept consistent where possible. The majority of notation has been confined to the Roman form, although Greek symbols have been used where such use is common in the literature. While some special uses are described more fully where they occur, a collection of the common notation is presented below. The majority of the terms may occur with associated subscripts.

a,b	any constants
c	sign and flatness of quadric terms
d	spatial distance in the horizontal plane between two points
e	error at a point
f	denotes local function
f()	denotes 'some function'
g(), h()	denote specific orthogonal functions
i, j	counts
K	Fisher's dispersion factor
n	number of cases, total number, selected sample size contributing to an interpolation function
p()	used to define specific spline functions
P()	sequence of orthogonal polynomials
q	coefficient of correlation
r	weighting ratio in interpolation
R	standardised vector strength
R ($r_1, r_2, \ldots r_n$)	chain in chain correlation
S ($s_1, s_2, \ldots s_n$)	chain in chain correlation
Rl	total vector strength
s	average search radius in interpolation
S()	denotes general spline function formula
t	tan (slope)
V	terrain variability measure
w	weighting factor in interpolation
w()	denotes a given positive function used in evolving orthogonal functions
W	wavelength of terrain
x,y,z	mutually perpendicular axes and their resultant spatial coordinates
α, β	any angles - see specific cases
Δ	increment or small amount

CHAPTER 1

INTRODUCTION

1 INTRODUCTION

1.1 OUTLINE

1.1.1 Subject

In 1965, Tobler stated:

> "...there are now several computer programs available which will draw isarithms when given observational information at only scattered locations. Of the several dozen such programs now in existence, however, no two seem to use the same mathematical logic. One wonders which will give the more correct maps. Almost everyone claims that his procedure is an improvement (in accuracy as well as speed) over the manual methods. Objective tests have not been published, but the majority of users seem satisfied" (Tobler, 1965, 36).

The most efficient method of comprehending data involves the medium of graphics. Where spatial data exist, most earth and some social scientists require the production of a map, more particularly an isarithmic map, to display their data. Such data are increasingly being stored in a computer-accessible form as a pre-requisite to complex data processing. The scientist is thus faced with a predicament of which is the optimum method of isarithm interpolation for his data. Manual interpolation is tedious and laborious, and yet automated methods inherently involve some loss of control over the process, and its inherent surface interpretation, to the computer. Even after the automated process is chosen, the

scientist is faced with a dilemma. Nearly two decades after Tobler first claimed the existence of 'several dozen' programs, there are now hundreds of algorithms and packages, each of which is claimed to be superior to other computer programs and to manual processes. Yet it is known that,

> "Many commonly used computer algorithms can produce non-sensical results" (Unwin, 1974, 278).

A research problem - the assessment of isarithm interpolation procedures involved in the production of isarithms - therefore exists and requires urgent solution for the benefit of other researchers. Yet, given that there are many scientists and researchers who possess differing data sets, how may they be directed towards realising the optimal isarithmic product from their data? This thesis attempts to provide appropriate guidelines.

In any research problem, the researcher must match the solution to the problem. Thus, since this particular problem has an empirical nature, it requires an empirical solution. Additionally, it may be considered a bivariate problem. The data sets may vary quite considerably and, to a great extent, their statistical properties are unknown. The method of isarithm interpolation may also vary, although to a lesser extent, since through integration of the computer into the research environment and the subsequent widespread exchange of software, one researcher may access the same packages as his colleague working in another environment. The empirical investigation must therefore combine an evaluation of the accuracy of the generally available isarithm interpolation software, with a predictive study involving an examination of

the power of various data descriptors in order to direct the user towards the optimum software for a particular data set.

A researcher logically applies expertise gained from his own discipline and interests to provide a solution to a problem. However, it is also considered that, within this research topic, the topographic sciences - and more precisely photogrammetry - provide the only method of fully investigating the spatial interpolation problem: they are focussed upon observable continuous surfaces, capable of measurement to any desired degree of accuracy and surfaces of widely varying characteristics may be studied. The problems of isarithmic interpolation of other surfaces, particularly those in other disciplines, may however be inferred from this examination and from discipline-specific information.

1.1.2 Aims

Study of the literature suggests that final isarithm accuracy varies with the nature of the discipline involved, the characteristics of the surfaces and the data sets used and, perhaps most importantly, the nature of the interpolation method utilised. The aims of this thesis are therefore as follows:

1. To evaluate the accuracy of some of the major isarithm interpolation packages and methods used in the academic environment (Chs.5, 6, 7).

2. To investigate the inter-relationships between the various methods of random-to-grid and isarithm interpolation (Ch. 2). This allows the user to make some estimates of the possible performance of methods not investigated herein.

3. To examine empirically the influence of data and, specifically, of surface characteristics - especially geomorphometry - in predicting the accuracies of different methods of random-to-grid and isarithm interpolation (Chs. 3, 4, 6, 7).

4. To assess briefly the conclusions of this thesis in the context of other disciplines and their data restrictions and general requirements within digital modelling (Chs. 2, 3).

1.1.3 Structure

The chapter structure of the thesis is as set out below:

Chapter 2 provides a general examination of the nature and uses of random-to-grid and isarithm interpolation in digital modelling. It examines the multi-disciplinary nature of the users and their uses of interpolation, stressing the

importance of establishing the regular grid. It outlines the methods of interpolation commonly utilised to produce such a grid and the associated isarithms, and subsequently examines the interpolation programs and methods evaluated within this research (2.5).

Chapter 3 considers the benefits of utilising photogrammetrically-derived topographic data within this research and closely examines the nature of the data. Additionally, by discussing commonly used methods of data classification, it relates the data used in this study to those used in the 'real world'.

Chapter 4 examines the descriptive power of various surface- and data-characteristics descriptors and, by determining their values for the various data sets used, demonstrates their suitability and consistency as accurate terrain descriptors.

Chapter 5 investigates the assessment of map accuracy and describes in detail the methods to be used within the research.

Chapters 6 and 7 assess the accuracy of the methods of interpolation examined and attempt to associate the errors generated with the surface and data characteristics. Chapter 6 is concerned with random-to-grid interpolation, while Chapter 7 examines the isarithm interpolation stage. Finally, Chapter 8 presents the conclusion of the thesis.

Appendix 1 describes the computer hardware and more specifically the software used within the research.

Presentation follows standard convention. Chapters, sections and subsections are labelled numerically using a common hierarchical notation. Figures and tables are separately distinguished and labelled numerically within each chapter.

Algebraic notation is intended to be consistent where possible and is defined when first used or changed. An overall summary of notation may be found at the beginning of this thesis.

1.2 OPERATIONAL DEFINITIONS

Several expressions, often considerably abused in the literature to the extent of ambiguity, must be defined for clarity within the thesis.

1.2.1 Isarithms

The terms 'Isarithm', 'Isoline', 'Isometric line', 'Isopleth', 'Contour' and all the Iso-terms are widely used and interchanged, resulting in variation and confusion as to their meanings. For the purposes of this thesis their use will be based on the definitions proposed by Robinson.

> "Isolines are the sets of lines (usually on a map) that show by their absolute and relative positions the locations and gradients within a set of numbers,...there has been an immense number of names coined for these lines, but in English language areas isarithm or isoline is generally employed as the generic form to include all such lines" (Robinson, 1971, 49).

Isarithms (or isolines) may be subdivided into two classes. One class, designated isometric lines ,

> "portray numerical distributions which may be termed

> basic, elemental geographic facts. The values they represent actually can exist at points on the Earth" (Robinson, 1971, 49).

Examples of isometric lines include elevation, air temperature, depth, atmospheric pressure. The other class, termed isopleth lines, portray the distribution of relative values. These are higher-order, more complex geographical concepts or abstractions which are a function of one element and space, for example density, spacing, ratios, correlations. Consequently, isopleth lines cannot exist at an exact point as they directly or indirectly involve an areal factor in their derivation.

In addition to the four general terms - isoline, isarithm, isometric lines and isopleth - there are numerous terms used in precise disciplines. For example contour - which is equivalent to the isohypse - is normally used solely to describe elevation. Similarly, isotherm is used for temperature distribution, isobar for pressure and isobath for underwater depths etc.

The term isarithm will be most used in this thesis since it is the general multidisciplinary term for lines which show geographical distribution of a specified value of a continuous variable.

1.2.2 Digital Modelling

The terms digital terrain model, digital elevation model, digital ground model and digital height model are frequently

used with little regard to the meaning of the intermediate descriptor used i.e. terrain, elevation, ground: considerable variation in meaning may also be observed in the literature. The following definitions are used throughout this thesis.

'Elevation' and 'height' have essentially the same definition, namely, the height or distance above a datum. Thus digital elevation model (DEM) or digital height model (DHM) are similar terms for expressing a collection or set of discrete heights above a datum. The elements may have any geographical distribution - points, lines, polygons, grids - with implicit or explicit connections.

The definition of 'ground' is that of a solid surface. A digital ground model (DGM) is thus similar to the DEM/DHM with the addition that there is some connection between the elements, which are no longer discrete. This connection generally takes the format of an inherent interpolation technique which may be utilised to regenerate any point on the surface.

The concept of terrain is more complex. Terrain not only includes the ground, but descriptive features on and near the ground. A digital terrain model (DTM) therefore uses a DGM as a reference to overlay other surfaces, for example vegetation, population and geology. The minimum requirement is a DGM, although the number of additional feature surfaces used is potentially infinite.

Finally, digital model is the general term for any surface. Like the previous model definitions, this may include any

point distribution: such surfaces may arise from any discipline.

1.2.3 Interpolation

A major aspect of isarithmic mapping is interpolation. The term interpolation is here defined to mean the process of estimating unknown data from known values within their environment. Thus the interpolation method is a procedure commencing with known data and ending with the establishment of a definitive value where previously unknown; the interpolation function is the mathematical formula used within the interpolation method.

1.3 SUMMARY

It has been stated earlier that a "researcher logically applies expertise gained from his own discipline and interests to provide a solution to a problem". This research has been undertaken in an environment of computer mapping. It is an empirical study which utilises some statistical and quantitative techniques; it is not a theoretical statistical study. From the outset the topic has been considered through the eyes of the user. Therefore, while some terms have been used which have a precise statistical meaning, such as 'parameter' and 'significance', in general these expressions should be considered in the user's context.

Additionally, many users may consider the thesis as a "Which ?"-type report. Results from the research will show (see later chapters) that there are no simple solutions to the problem of choosing interpolation procedures but, as with many research problems, the answer is a complex weave of 'ifs' and 'buts' and therefore must be considered in a broad context.

CHAPTER 2

INTERPOLATION

2 INTERPOLATION

2.1 INTRODUCTION

Computer cartography has been in existence for over two decades. Its humble beginnings as line-printer mosaics, generated from the only currently available hardware, have developed into precision graphics utilising high-technology hardware capable of surpassing manual cartographic quality. In addition, the computer solution is flexible and increasingly accessible and therefore the range of users and their applications has increased, developing the simple computer cartography into digital modelling.

An important aspect of this work is the processing of the initial, observed data into a convenient form for manipulation and final map production. Most processing, especially of the third dimension (assuming the first two dimensions represent the plane of the map sheet), involves interpolation. Interpolation is required as, due to some reason, a point value is required which was not observed. Just as data sampling may be considered a form of data compaction, so interpolation can be considered a form of data reconstruction.

The interpolation of isarithms from random spatial data in digital modelling is invariably subdivided into a two-stage procedure (Figure 2.1). This involves the establishment of some intermediate grid (2.4.1 and 2.4.2) through which isarithms are then threaded (2.4.3).

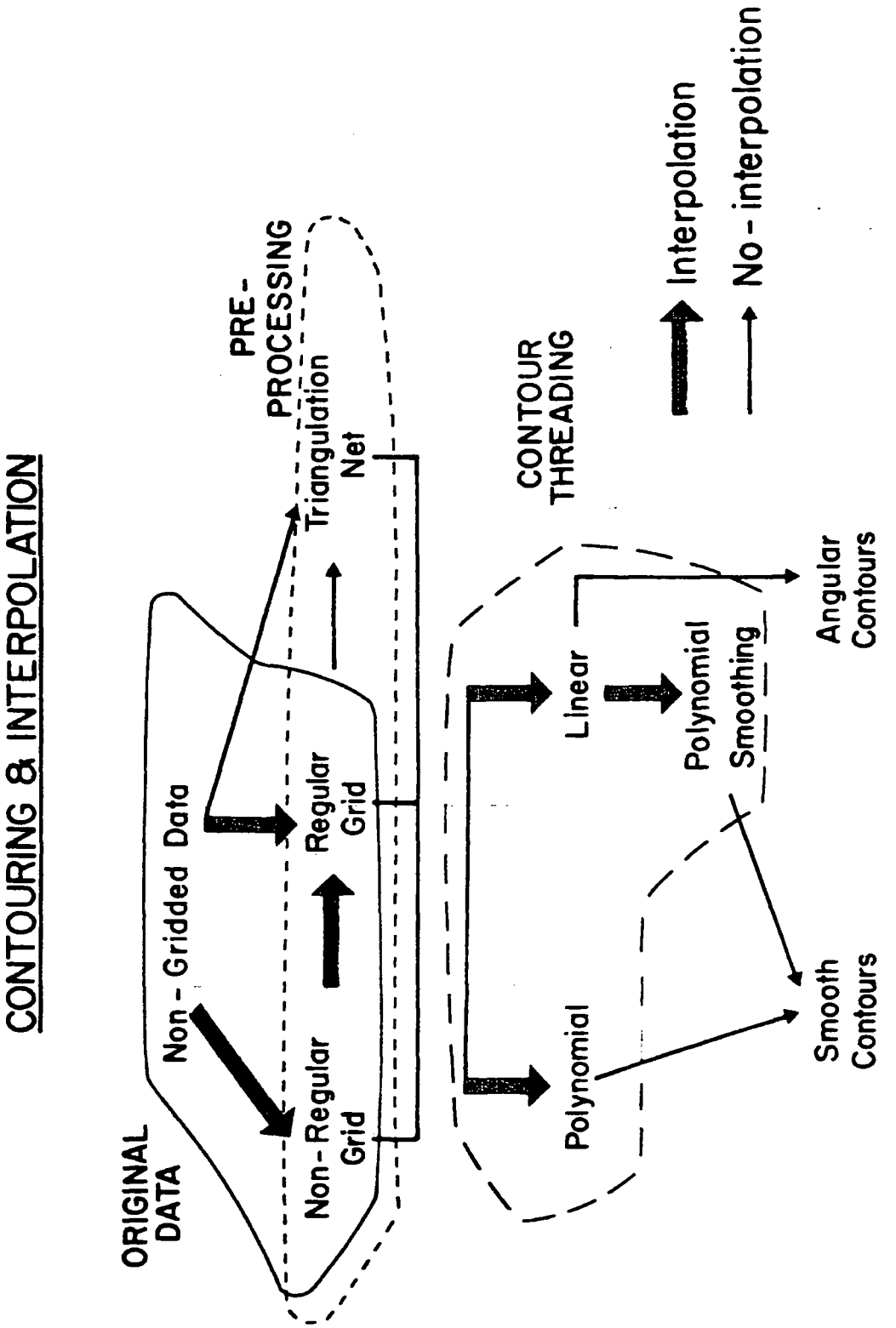


Figure 2.1 Contouring and interpolation

The grid may be of three forms. Each form may be used as the starting point for generating isarithms and often one may be derived from another (Figure 2.1). The most popular method involves the interpolation of a regular square (or rectangular) grid (see 2.4.1 and Figure 2.2a/b). This is the most complex and varied aspect and its discussion will therefore represent the major part of this chapter and thesis, much of it having relevance to the actual interpolation of the isarithms from the grid. Equally, a non-regular rectangular grid may be interpolated from the data. This is based on the theory that the grid resolution should vary with the nuances of the terrain. It involves similar considerations to the derivation of the regular grid except that the resultant cells vary in size. In practice it often represents an intermediate step to the generation of the more flexible regular grid.

The grid may also be triangular (see 2.4.2). Often this may be regular and defined from a regular grid either by interpolation or simple re-arrangement (see 2.4.2.1). Either a central point is interpolated and the grid sub-divided into 4 triangles, or the cell is diagonally bisected into two triangles. Alternatively, an irregular grid is often derived without the use of interpolation by applying an algorithm to the original point data (see 2.4.2.2 and Figure 2.2a/c). There are many such algorithms, but most either seek to keep triangles equiangular or attempt to derive short-sided triangles.

The resultant rectangular grid or triangulation net (Figures 2.2b and c) represents an important concept in digital

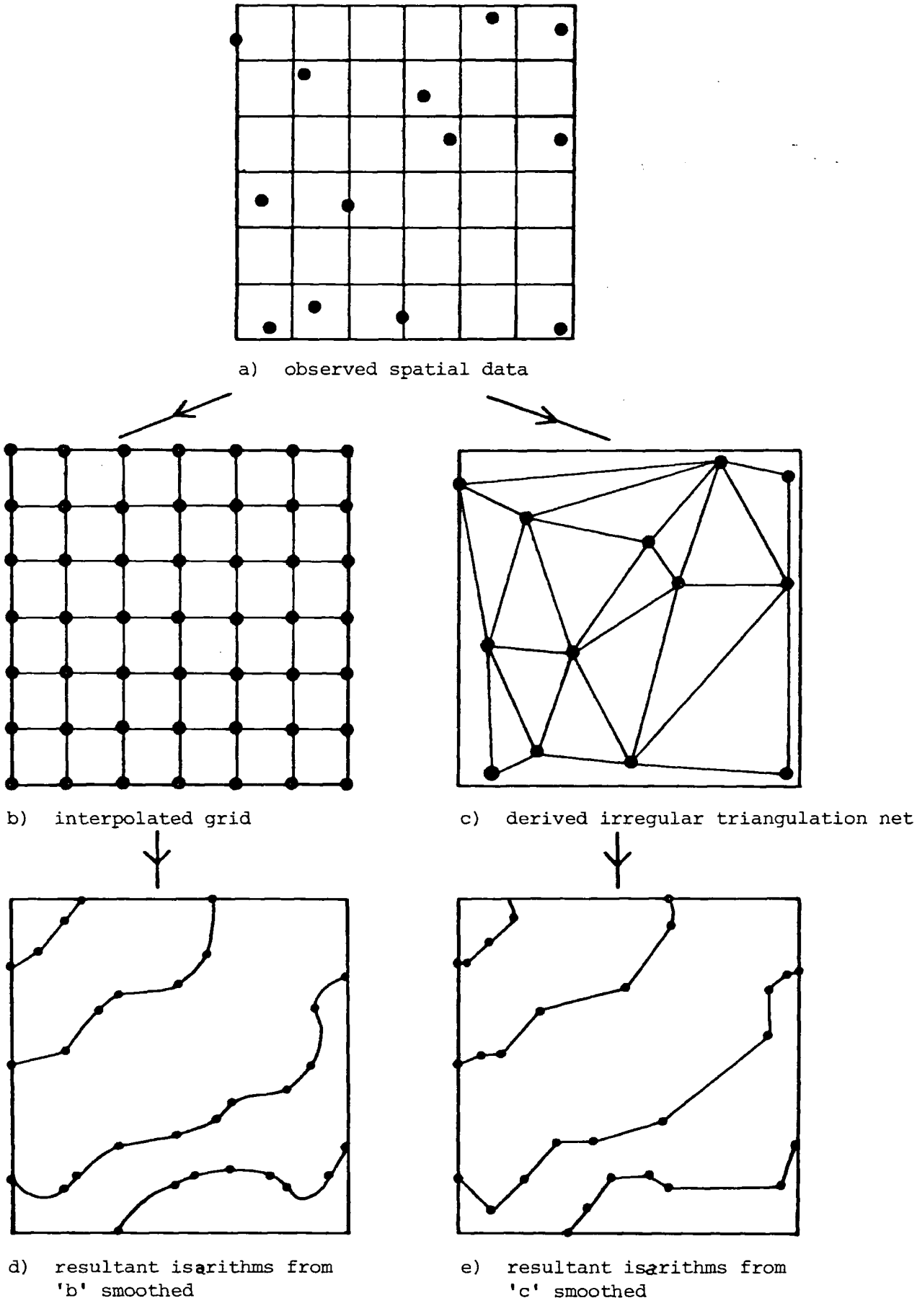


Figure 2.2 Graphical summary of the use of interpolation in isarithm production in digital modelling.

modelling. This will be demonstrated more clearly in section 2.3. Essentially, these digital models (grid or net) may be used to produce a wide range of varying products. The kernel task in digital modelling has been performed, and therefore the digital models are generally saved as a new starting point for any future production or use. The raw data are henceforth considered of only secondary importance.

In isarithm generation (see 2.4.3), isarithms will be threaded through the model using some simple interpolation process. To produce smooth isarithms directly (see Figure 2.2d), a form of patchwise interpolation may be utilised involving threading isarithms through temporarily interpolated, finer resolution grids. Alternatively, a simple linear interpolation only may be used in the threading process, producing angular isarithms (see Figure 2.2e), which may be subsequently smoothed using a piecewise polynomial.

In conclusion, this chapter involves a discussion of the interpolation process involved in isarithm interpolation. Use of this process is not restricted to one discipline and therefore it is important to consider it in the context of other disciplines, especially with regard to their data requirements (2.2): this may also allow other users to relate the conclusions drawn in this thesis to their disciplines. The establishment of a grid or net from observed data has already been stressed. In examining the uses of interpolation in digital modelling (2.3), the importance of the random-to-grid and triangulation methods is illustrated as a basis for any application. The interpolation methods commonly used in

isarithm generation are examined in greater depth in section 2.4. This, in conjunction with the examination of methods investigated in this research (section 2.5), allows the user to relate research results to other methods not evaluated.

2.2 USERS OF THE INTERPOLATION PROCEDURES IN DIGITAL MODELLING

Most of the discussion within this thesis is concerned with topographic modelling and its associated data. While such modelling has provided the initial impetus to the development of many of the computer applications, users in many other non-topographic/engineering disciplines have realised the potential and thus use of the same procedures has become more universal.

Topographic mapping involves modelling a wide variety of surfaces. The majority of applications involve modelling at a large topographic scale and are associated with the original development of DTMs by Miller and Laflame (1958) (e.g. Maxwell, 1970; Craine, Houlton and Malcomson, 1974; Keir, 1976 and Assmus and Stanger, 1978). These users employed a wide variety of data distributions (points, lines and grids) with surfaces often highly autocorrelated between sharp discontinuities. At the other extreme, small topographic scale modelling employs more regular (perhaps gridded) data (e.g. Zarzycki, Harris and Linders, 1975; Bethel, Crawley, Shepphird and Hussain, 1978 and Dutton, 1982): the surfaces thus described may be moderately 'rough' as a result of the small scale and coarser resolution data. Medium scale users employ a wide variety of data types (generally surface feature-specific and often

lineated) and surfaces (very often rough to relatively highly autocorrelated); these users are generally concerned with data banking of some description (Bell, 1977 and Allam, 1978). Interpolation is utilised within this broad spectrum of users primarily to establish regular grids as an intermediate stage to the generation (and therefore interpolation) of the full spectrum of applications discussed later.

Bathymetric and oceanographic mapping generally deal with surfaces which are essentially similar to those in topographic mapping. However, as a consequence of the medium-to-small scale requirements of most such mapping, a wide range of surface roughnesses are encountered. In addition, large scale mapping of the environs of oil rigs and pipelines utilising sonar digital devices is being increasingly attempted, for example within the recent North Sea exploration. Some of the earliest computer-produced isobathic charts were constructed by Swindle and van Andel (1969) of the mid-Atlantic ridge. They claimed these to be more efficiently produced than conventionally contoured charts, yielding an acceptable representation of sea-floor topography free of interpretative bias. More recently, several other workers have demonstrated the automated contouring aspect of chart production (e.g. Boyle, 1971 - at the Canadian Hydrographic Service; Richardson, 1975 - at the White Fish Authority; Slootweg, 1978 of an area across the mid-Atlantic ridge and Peucker, 1980). Non-contour products produced include the early computer production of anaglyph stereoscopic bathymetric maps (Lang, 1969), the later production by the ECU of perspective plots of tracked data in the Thames estuary, utilising the SACM package

(Thorpe, 1971) and the general production of stereoscopic maps and perspectives (Peucker, 1980).

Naturally, geographers have been major users of some of the techniques. Bunge suggested as early as 1962 that only 'pure hideboundness' keeps cartography from applying isarithmic techniques to a wide range of geographical phenomena. Many geographical papers now include computer-generated isarithm maps or block diagrams. Many of the early isarithm maps were derived from trend surfaces fitted to irregular data (e.g. Chorley and Haggett, 1965 and Robinson, 1972). However, geographical usage, which often tends to involve inherently smooth surfaces, also includes many computational applications (e.g. Monmonier, 1971 - conducting research into orchard location; Monmonier, Pfaltz and Rosenfeld, 1966 for the derivation of SAMP; Speakman, 1980 - predicting noise exposure; Collins and Moon, 1981 and Spanner, 1983 - modelling soil erosion).

Geological workers, especially Davis (1973), have always been major users of digital modelling techniques (Moore and Simpson, 1982 and Watson, 1982). Work includes the evaluation of different contouring methods by Rhind (1971 and 1972), Walden (1972), Barrett and Rhind (1974) and Bolondi, Rocca and Zanoleti (1976). Some of these comparisons include data with sharp faults dissecting highly correlated inter-fault regions. Additionally, Cubitt and Celenk (1976) and Sampson (1978) - through the development of the SURFACE II GRAPHICS package - have demonstrated the growing application of stereoscopic and perspective plots to geological mapping. Geological data are

generally sparse, poorly distributed and relate to a constant time horizon or to a lithostratigraphic unit. However, as a result of the cost of obtaining the data, the data points (bore holes and wells) are generally surface-specific, their location based on some prior knowledge of the surface. Surface roughness varies considerably with rock type and, if exposed at the ground surface, with the degree and type of erosion suffered by the parent material.

Geophysical digital mapping often utilises lineated data (in the form of 2D seismic data) with approximately gridded data (in the form of 3D seismic data) becoming increasingly popular more recently (McCullagh, 1983). Surface smoothness can vary as with geological data, although much of the literature concerns itself with highly correlated surfaces (Reilly, 1981). Huijbregts and Matheron (1970), Bhattacharyya and Ross (1972) and Braile (1978) have investigated particular algorithms in relation to such data. There appear to be few non-contouring applications of digital modelling in geophysical computer mapping.

Within mining engineering, much of the data are clustered and poorly distributed. Some of the earliest studies were undertaken by Krige (1966) in his examination of the advantages of (what is now referred to as) kriging in borehole valuations of developed and undeveloped ore reserves in South Africa. Computer isarithmic and interpolation work has also been performed by Agterberg (1969), Gold, Charlesworth and Kilby (1981) and Royle, Clausen and Frederiksen (1981). Additionally, Gimbert and Cubitt (1973), and Bethel, Grawley,

Shepphird and Hussain (1978) have performed volume estimation of coal heaps from digital models.

The definition of meteorological surfaces, which are highly spatially correlated and constantly varying in time, causes severe data acquisition problems. In addition, a fast, efficient production of results is often demanded by the user. A further complication is that meteorological data generally consist of sparse point data which may be poorly distributed in spatial terms. Two summaries of computer isobar production for raw, internationally provided data are given by Menmuir (1974) of the British Meteorological Office, and Gordon (1981) of the New Zealand Meteorological Service. Additionally, the successful implementation of the ubiquitous multiquadric analysis is comprehensively demonstrated by Shaw and Lynn (1972) and Lynn (1975) to rainfall evaluation prediction. More interestingly Adams (1969) demonstrated the ability of anaglyph stereoscopic weather charts to overlay the sea level pressure and heights of the 700 millibar layer around the Continental Shelf and Speakman (1980) has considered temperature and humidity changes around an airport.

One of the earliest grid-to-isarithm programs was established for chemical data sets. Dayhoff (1963), describing a program which has subsequently been the basis for many other packages (eg. GINOSURF), showed that the computer could produce isarithm maps of electron density and protein planes more efficiently than the laborious conventional method. Walden (1972) has subsequently demonstrated computer-generated isarithms for analysing variation in yield from a reaction

over a plane of changing pressure and temperature, and Cubitt and Celenk (1976) have used stereograms to map the chemical concentrations of streams. Such chemical data sets are either irregularly and poorly distributed, low autocorrelated data collected in the field (Kane, Begovich, Butz and Myers, 1982), or else are regular grids of highly correlated data observed in the laboratory.

Medical surfaces are similar to the laboratory derived chemical surfaces. Renner, Bahr and Compaan (1980) describe isarithmic problems in radiation therapy, the data being photogrammetrically produced of radiation from a Co-60 source. van Oosterom (1978), utilising isarithms and perspective views, has examined similar problems in dose distributions in radiotherapy and radiology and Cook, Dwyer, Batnitzky and Lee (1983) have utilised perspectives and volumetrics in diagnostic applications.

Mathematical surfaces, although apparently complex from their formulae, are in reality very smooth. Morrison (1974a, 1974b) and Walden (1972), for instance, both used mathematical surfaces in the study of interpolation, Morrison also producing block diagrams. Both authors have particularly chosen inherently smooth surfaces, Walden using a surface of the formula,

$$z = 10000 - (x - 61) - (y - 61),$$

while Morrison used the selection of surfaces found in Figure 2.3.

The military, especially in aeronautical and radar applications, is responsible for much research into the

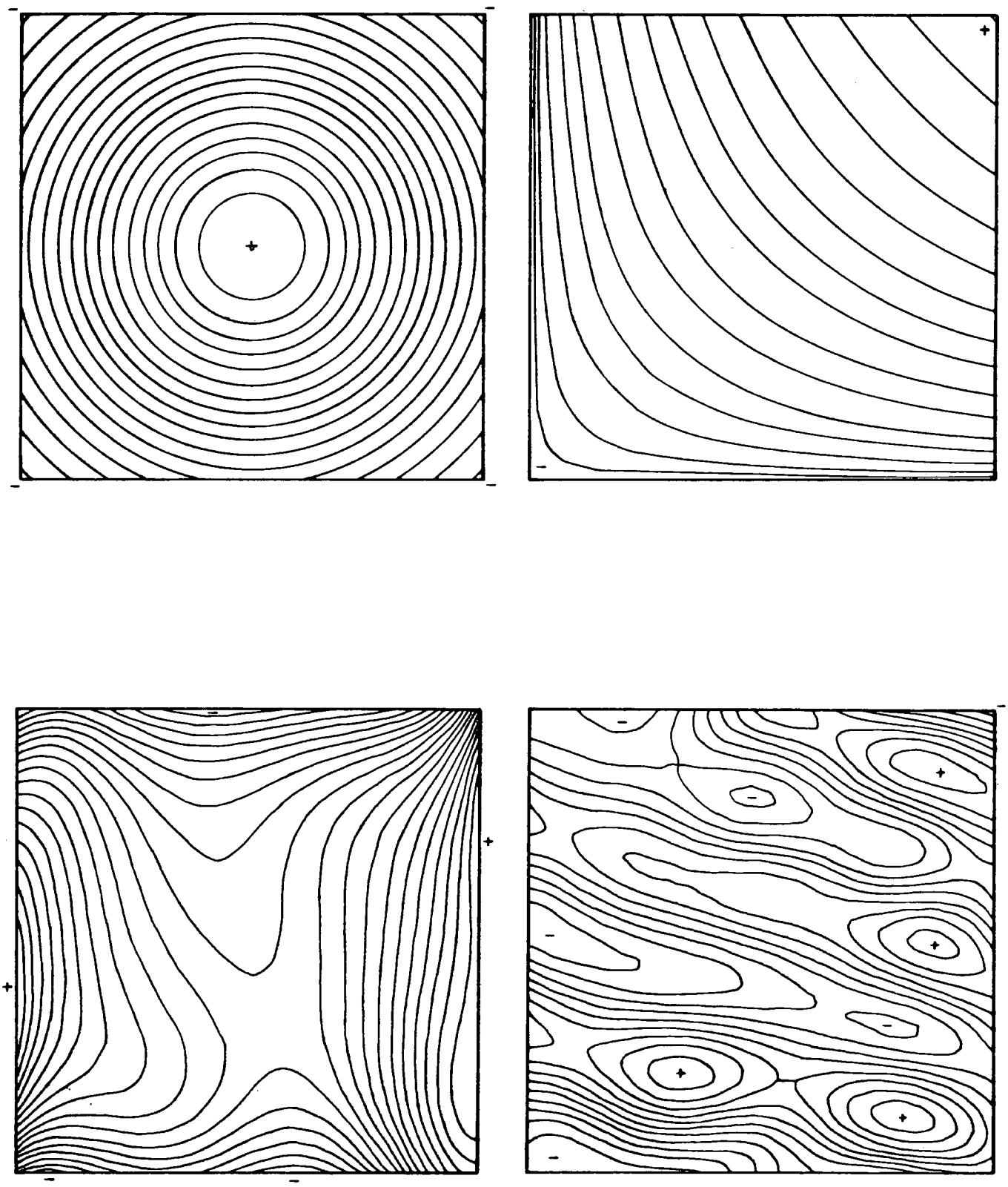


Figure 2.3 Selection of mathematical surfaces used by Morrison to investigate interpolation error (after Morrison, 1970)

computational and real-time aspects of digital modelling. The use of digital mapping in aeronautical work and planning involves some of the recent and futurist applications. In aeronautical work, flight simulation and digital animation are major users of perspectives and visibility studies. Ottoson (1978) reported on the use of the Swedish elevation data base for flight simulation systems and radar visibility studies, Shoup (1979) discussed the use of 'Superpaint' in the digital animation of some of NASA's missions and Jancaitis and Moore (1978) present a graphical contribution on the use of near real-time perspectives at the US Topographic Engineering Laboratories. Rhind (1975) also noted the isarithmic mapping of signal strengths from radio and TV transmitters and wave distortion in electronic circuits. Stein (1959) was involved in perhaps the earliest study of Line Of Sight (LOS) problems and more recently Ackermann and Crombie (1975), Frye (1978), Ottoson (1978) and Olender (1980) have also contributed to the subject. In these examples, the data are usually regularly distributed in space. Excluding terrain, surfaces in military use (e.g. electromagnetic) are usually disjointed, with highly correlated patches between sharp breaks, and are thus superficially similar to many geological surfaces.

It is obvious from the previous discussion that digital modelling has become a tool for many surface-related applications in many disciplines over the last decade. While most individual users have a more limited variety of data samples, in general a full range of data types representing a wide cross-section of surfaces are used within interpolation procedures in digital modelling. These observed data are

generally processed to create a more flexible data format (regular grid/triangulation net), which forms a basis for the secondary processing - generating the specific product desired.

2.3 USES OF THE INTERPOLATION PROCEDURES IN DIGITAL MODELLING

2.3.1 Introduction

There is no doubt that the observation and logging of data in a digital environment is often a more time- and/or cost-consuming method of data collection than conventional methods. While equipment such as automatic line following digitisers or automatic orthophotomappers do exist to speed up the actual logging process, the increased efficiency may usually be balanced by increased machine costs and/or pre-logging data preparation requirements. Additionally, the data must generally undergo some post-logging processing, although in production environments this may be kept to a minimum. The raison d´être for this increased effort and cost is that the data are now in a more flexible format and many more uses may be made of them.

Digital modelling, then, has evolved from the realisation that, having performed the major operation - creating the representation of a surface in the computer - this may be used to solve many problems directly, rather than indirectly through the graphics medium. Three types of applications may therefore now be discerned. Graphical output includes isarithm

maps, block diagrams, profiles, stereoscopic views and other specialised methods of relief representation. Secondly, applications based on purely numerical products include calculation of profiles, areas and volumes and line-of-sight evaluations. Finally, engineering and structural design represents a major growth application. Beyond these application types, a considerable amount of effort is being spent on data organisation in the form of a data base or data bank, which is used as the basis for the generation of the other three products.

These applications are directly related in that they all employ interpolation, either as a pre-requisite to computation or during generation of the products. In order to increase flexibility and to facilitate subsequent processing, all users of spatial data require data to be in an organised form. This generally takes the form of a regular grid (necessitating random-to-grid interpolation) or occasionally a triangulation net (requiring some form of data reorganisation). Additionally, all subsequent processes require interpolation (e.g. grid-to-isarithm) of the intermediate mesh (grid or triangulation). These processes must therefore be considered in the context of how they employ interpolation.

2.3.2 Graphical Products

There are three commonly generated graphical products which may be classified in terms of the way the digital model is viewed. If the model is viewed from a vertical infinity, the resultant product is planimetrically correct. Alternatively,

if the model is viewed from a non-vertical position, some form of block diagram is generated. A profile may be considered a special form of block diagram restricted to two dimensions. Finally, stereoscopic views may be produced of either of the previous two products.

2.3.2.1 Plannimetrically Correct Products.

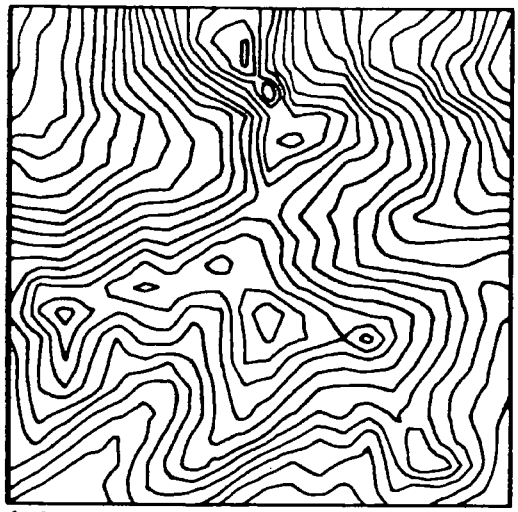
The generation of planimetrically correct products is of special significance in the display of data, as it allows the overlaying of maps and other data in their correct geographical position. It is therefore the most popular method of presentation and includes both raster and vector derived products.

2.3.2.1.1 Isarithmic Mapping.

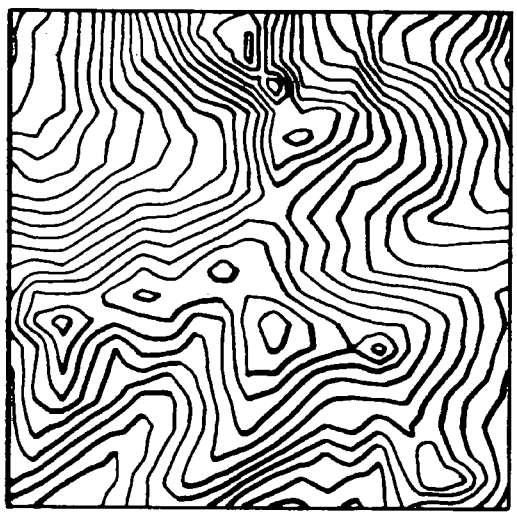
Three forms of planimetrically correct isarithm maps are frequently computer-produced.

The standard isarithmic map (Figure 2.4a) is the most popular graphic having the advantages of simplicity of production, computation and extraction of information. This is formed by threading the isarithms (by interpolation) through a regular grid of data or triangulation net.

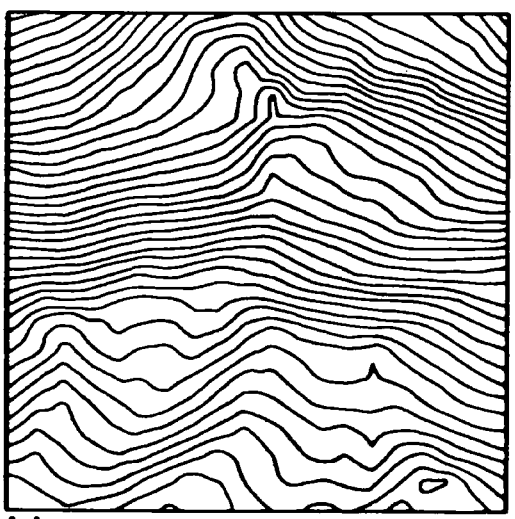
Relief isarithms (Figure 2.4b, and discussed for example by Peucker, 1972; Sprunt, 1975 and Yoeli, 1983) simulate the illumination and shadow effect of isarithms on an illuminated surface. Generally, a regular grid of data is processed to generate a numerical value of pseudo-brightness at each grid



(a) Standard Isarithms



(b) Relief Isarithms



(c) Oblique Isarithms

Figure 2.4 Various forms of isarithm maps (after Sprünt, 1975)

node. Isarithms are then interpolated through this secondary grid in the same way as for standard isarithms, but are plotted with varying line widths.

Orthogonal relief isarithms (Figure 2.4c) - frequently termed oblique, or more descriptively inclined isarithms - have the advantage that they

> "permit construction of a perspective-like view of the terrain that is planimetrically correct" (Robinson and Thrower, 1957).

The automated method for construction of these isarithms, discussed by Peucker, Tichenor and Rase (1975) and Sprunt (1975), is simpler to conceive than that for relief isarithms, as background shading and variable isarithm thicknesses are not required. Isarithms may be traced conventionally through a regular grid if it is first pseudo-inclined using the formula,

$$z_{new} = (x.\sin(\alpha).\sin(\beta))+(y.\cos(\alpha).\sin(\beta))-(z_{old}.\cos(\beta)),$$

where x,y = the spatial co-ordinates in the horizontal plane,

z_{old}, z_{new} = the original and inclined z values,

α = the azimuth rotation (i.e. around the z axis),

β = the altitude rotation (i.e. around the selected azimuth reference).

2.3.2.1.2 Hachures.

The generation of hachures (Figure 2.5a) depends on a systematic classification of the digital model into slope angles. Where a regular grid exists, slope may be easily

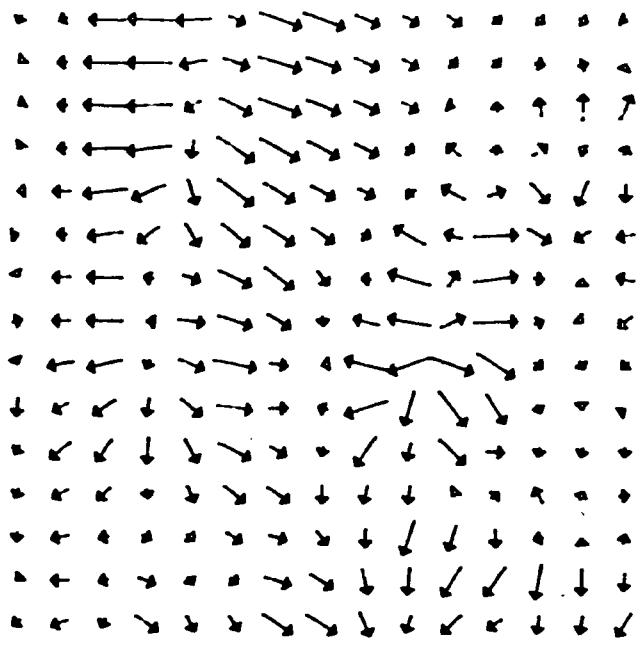


Figure 2.5a Hachure map of FORV data set (see Chapter 3)

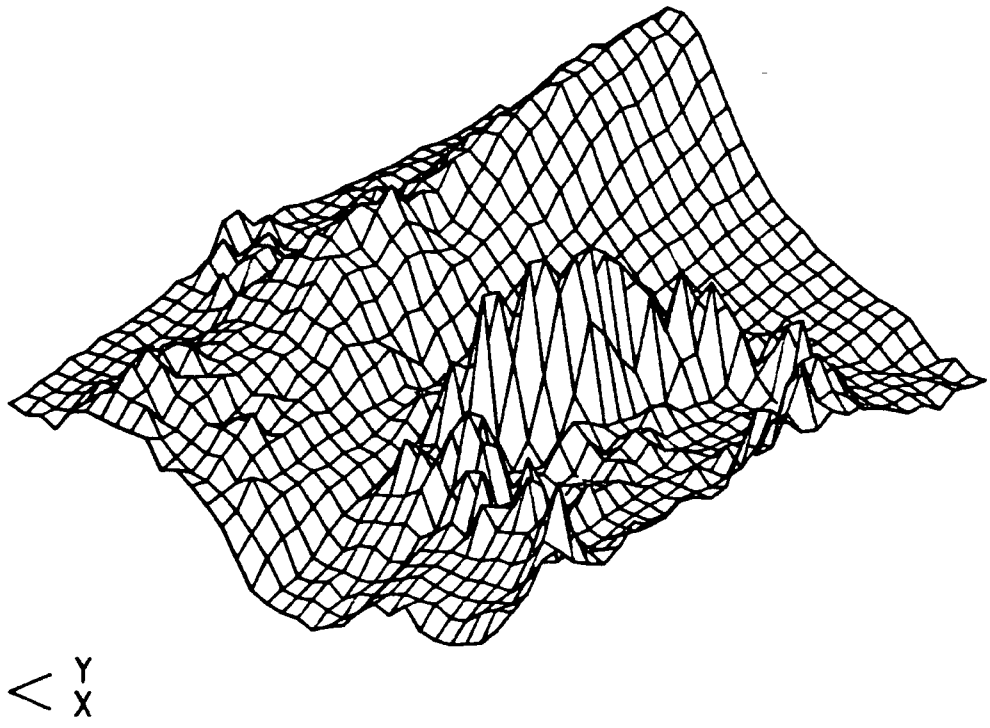


Figure 2.5b Block diagram of FORV data set (see Chapter 3)

evaluated at each node by considering the surrounding nodes. Once classified, the hachures may be drawn (eg. Peucker, 1972 and Evans, 1979), the lines usually only varying in direction and length to suit the automated draughting process. This method of relief representation is most suitable where vector graphics are used to display the map.

2.3.2.1.3 Hill-shading.

Where raster graphics are being used, slope and relief are better illustrated by hill-shading (see Yoeli, 1965; Brassel, 1974; Anda, 1976 and Dutton, 1982). This technique, which requires the prior existence of a regular grid, employs computation related to that for relief isarithms and hachures, involving consideration of gradient, aspect and light reflectance. Output is ideally performed utilising a high quality raster plotter, where the area can be systematically scanned, a varying density of exposure being methodically applied to each pixel. Successful results may only be achieved if substantial photo-reduction of the plot takes place or if the grid has a very fine resolution.

2.3.2.2 Block Diagrams.

> "The block diagram or perspective view of a surface is gaining popularity as are other mapping methods which provide an easier and quicker comprehension of the surface displayed" (Peucker, 1972, 43).

Block diagrams (Figure 2.5b) are being increasingly derived

of either contours (e.g. within Radian Corporation's CPS-1 - contour plotting software package) or, more popularly, a regular grid (Cook, Dwyer, Batnitzky and Lee, 1983). If the latter is input, subsequent profiles are orthogonal and may be parallel to the axes or follow the diagonals. Regardless of the method, an inherent problem which must be overcome is that of hidden line removal - the problem that nearer high points obscure lower distant points.

There are three main perspectives used in block diagrams. One point perspectives (Figure 2.6a) used in, for example, MOSS (Craine, Houlton and Malcomson, 1974) and CIIS (Applied Research Cambridge, 1979b) are analogous to standing at a point observing the area. Isometric projections (Figure 2.6b) are the most popular form of block diagram (Peucker, 1972). Used in SYMVU, PERSYS, GHOST and CPS-1, the basis of this method is that the viewing point is considered to be at infinity, thus the viewing rays are parallel. A cylindrical projection (Figure 2.6c) is analogous to actually standing in the model and viewing the complete environs (Peucker, 1972).

2.3.2.3 Stereoscopy.

Computer-simulated stereoscopy has previously had limited use, but is gaining popularity and can now be generated by several packages (e.g. GPCP and SURFACE II GRAPHICS). Published examples include Adams (1969), McCullagh and Sampson (1972), Cubitt and Celenk (1976) and Peucker (1980). Indeed McCullagh and Sampson 'recommend' a routine to generate a stereoplot as an 'optional extra' in any graphics packages in

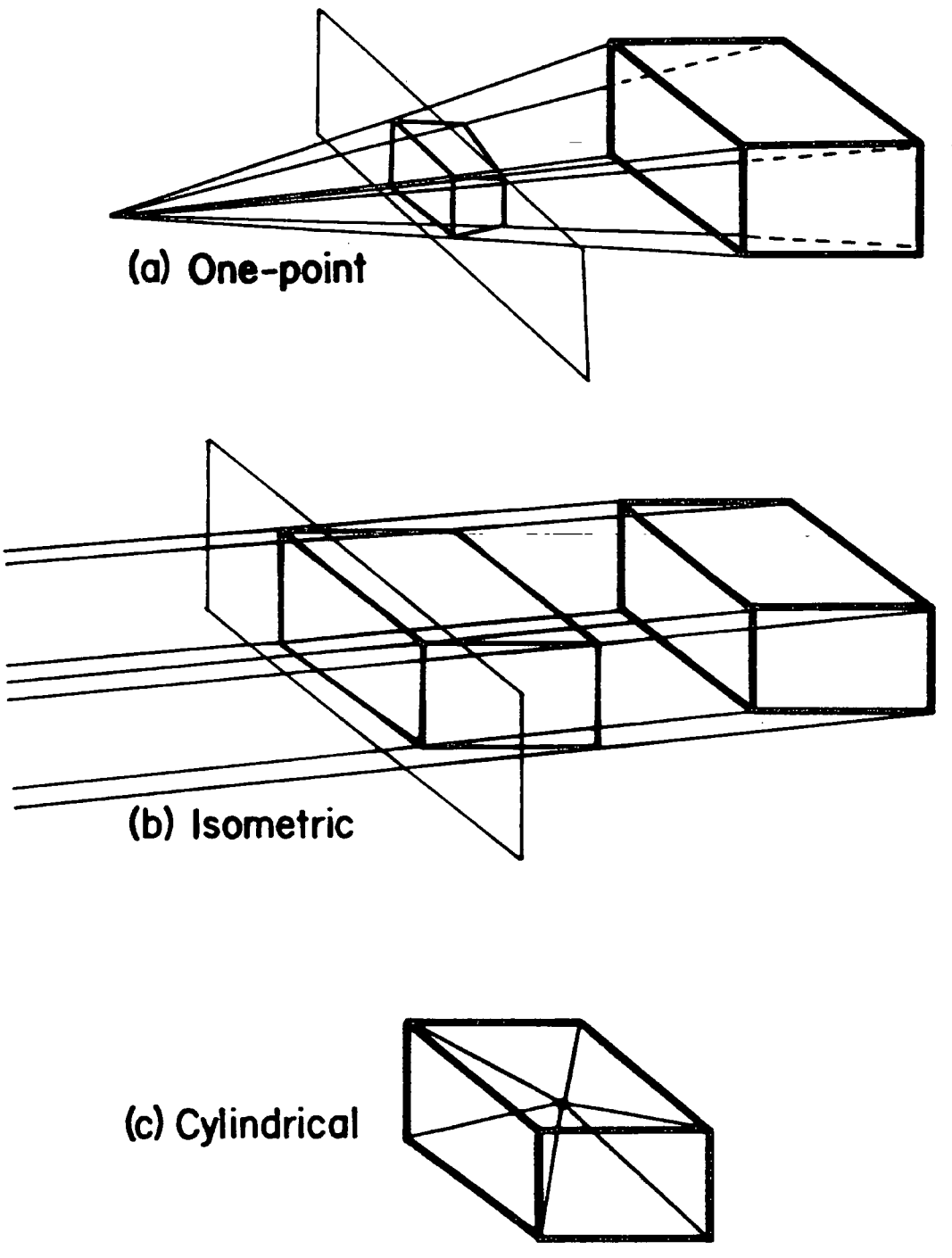


Figure 2.6 Types of perspectives used to generate block diagrams

academia.

The stereoscopic view may be generated of any graphical display, for instance block diagrams (Figure 2.7) - as in SURFACE II GRAPHICS (Sampson, 1978) - or isarithm maps (Figure 2.8), as in GPCP (Calcomp, 1973). In addition, the stereo display may be generated for viewing by anaglyph filters (Figure 2.8), pocket stereoscope (Figure 2.7) or mirror stereoscope. In all cases, the basic data and computation methods are similar to those in the monoscopic equivalents. The effect is achieved by applying a subsequent shift or rotation of the constituent plots and applying a separation (or parallax).

2.3.3 Numerical Applications

There are three main numerical applications of digital modelling. In the original development of digital terrain models (Miller and Laflame, 1958), the model was used for the solution of highway engineering problems. This still represents the major usage of large scale engineering models. At the other extreme of scales, military and radio applications involve evaluating line-of-sight problems.

2.3.3.1 Cross-sections, Areas And Volumes

Although applications in many disciplines (for example geomorphology and geology) involve the derivation of profiles, areas and volumes, the major users of this application are building and highway engineering. In the latter, the data generally consist of the co-ordinates of the proposed route

Figure 2.7 Stereoscopic block Diagram of FORV data set (see Chapter 3) for viewing with pocket stereoscope

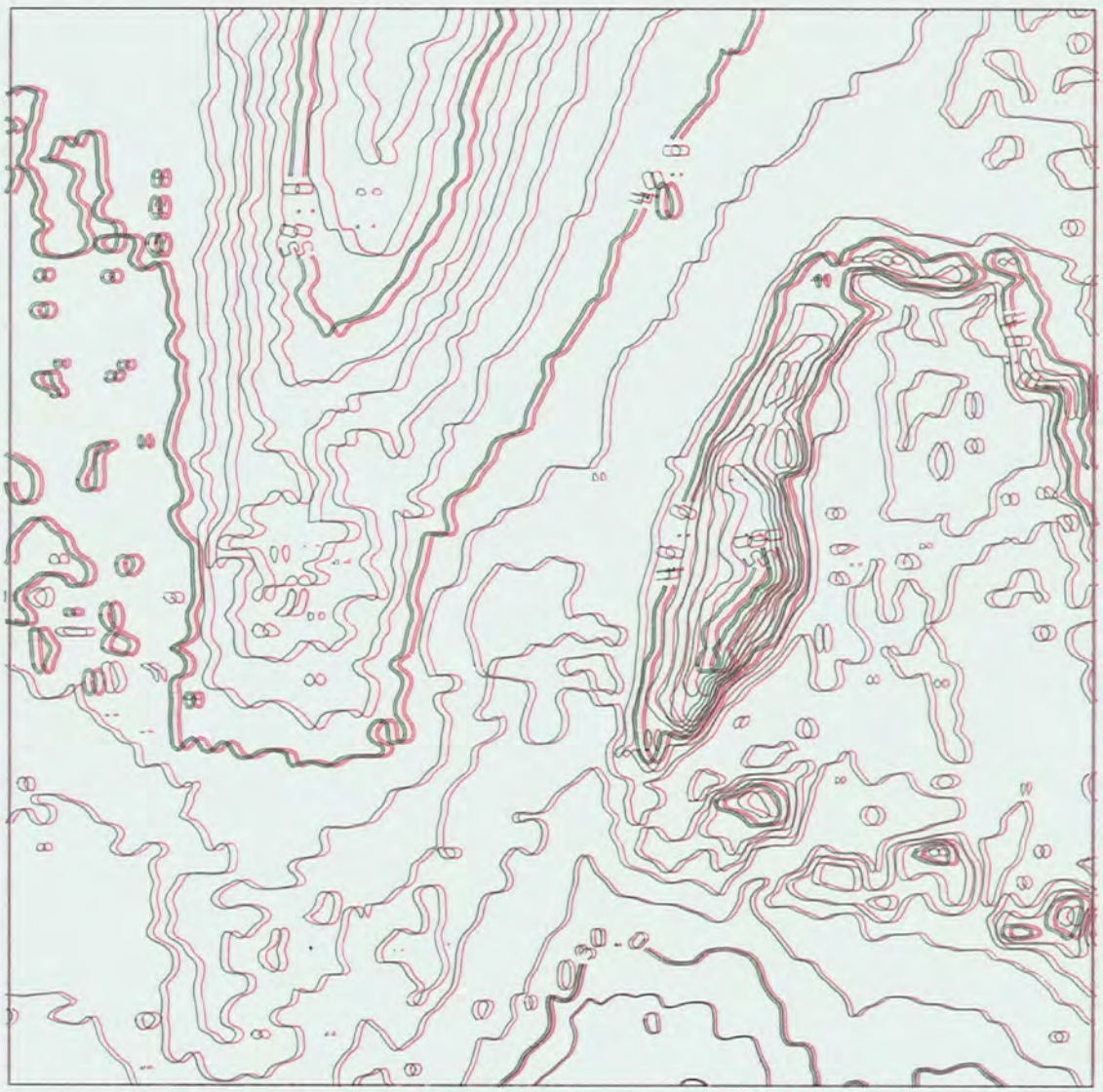


Figure 2.8 Stereoscopic isarithm map of FORV (see Chapter 3) for viewing with anaglyph

and a dense DGM which must be processed to form a reference base for subsequent computation. BIPS (Maxwell, 1970), SCOP, VIAK-AB and Nordisk ADB (Torlegard, 1972) and most methods discussed by Blaschke (1968), Keir (1969) and Maxwell (1970) interpolate a regular grid for this purpose. Additionally, many non-engineering programs, for example SAMP (Monmonier, Pfaltz and Rosenfeld, 1966), GPCP (Calcomp, 1973) and Cook, Dwyer, Batnitzky and Lee (1983), base their surface area computation on regular grids. However, CIIS (Applied Research Cambridge, 1979b) and MOSS (Malcomson, 1973; Craine, Houlton and Malcomson, 1974 and Institute of Civil Engineers, 1977) utilise a triangulation net. Some workers, for example Richardus (1976) and Nasca (1979), have shown that areas and volumes may be computed to a sufficiently high accuracy directly from contours.

Once the grid, triangulation net or contours are established, and the route (if a highway) or area of interest has been defined, the computation is trivial. Any additional extra information is evaluated by linear interpolation from the model, and the area/volume computation follows standard engineering formulae.

2.3.3.2 Line-of-sight (LOS).

The line-of-sight problem - the problem of inter-visibility - is mainly of concern to the military, those concerned with problems associated with the masking of radio waves and in civil engineering (e.g. Stein, 1959; Ackermann and Crombie, 1975; Frye, 1979 and Olender, 1980). The methodology for

solving LOS is based on the prior establishment of the two points of interest and the subsequent derivation of the connecting profile. This simple linear relationship may be developed into an areal relationship by considering several such profiles (e.g. in evaluating the area of visibility of a new chimney in civil engineering). While the earlier work of Stein and of Ackermann and Crombie utilised regular grids as a basis for establishing the visibility connecting profiles, Frye has more recently contrasted their success with the relative ease of use of triangulation nets. The surfaces being considered represent a wide range of surface roughnesses although large scale civil engineering projects have more recently provided surfaces with sharp discontinuities as a result of, for example, tower blocks, factories and other angular structures (Fullenwider, Lefever and Martin, 1981).

2.3.4 Design

The application of digital modelling in design (leading to CAD - Computer-Aided Design) encapsulates and integrates the previously discussed applications (Fullenwider, Lefever and Martin, 1981). The disciplines which utilise this technique increase daily and include design of highways, building sites and machinery (for example cars and smaller household items). Since the success of the design relates to the completeness of the computer model, a full variety of point, line, polygon and gridded data is often used. The multi-valued surface model is usually based on a regular grid or triangulation net. In road design, the majority of data are terrain line, contour or point data; this must therefore be pre-processed (e.g. as in

MOSS and BIPS). In the second stage of the design, where more specific and localised data are required, other data (e.g. land values and information about the existing nature of the road, fences and walls) may be added. Once the design is complete, isarithms, perspectives, volumes and many other previously discussed products are generated.

2.3.5 Data Base Or Data Bank

Associated with the design process and also with the ubiquitous information systems are the data bank and data base: in this, no single application is intended for the data and, indeed, the final uses may not be known when the data are collected and stored. Examples of these data storage systems are numerous (e.g. Bickmore, 1967; Schmidt, 1969; Irwin, 1971; Zarzycki, 1976; Harrison, 1978 and Spicer, 1979). To increase the speed of retrieval and the flexibility of the observed data, a reference surface, usually a grid, will be established for the whole area of interest. Efficiency obviously increases as more data are gridded-up using the same x,y reference co-ordinates. Such a procedure is frequently used in the oil industry where most data are gridded as soon as possible, thereby allowing stacking of zones and geological phenomena on a common reference.

2.3.6 Summary

It has been shown that, once having performed the time- and/or cost-consuming part of acquiring the data, a wide variety of applications are open to the user in a digital

modelling environment. While isarithm and perspective drawing are truly multi-disciplinary, other graphical products also have considerable potential outside topographic mapping. Many of the computational methods are primarily linked to specific requirements (e.g. LOS to military and electronic applications). However, it is conceivable that, with a wide-spread development of data banking and design, these applications may additionally become more universally utilised.

Isarithmic mapping may be the only present definitive multi-disciplinary application, but, before most other products may be generated, a regular grid or sometimes a triangulation net must be established from the raw input data. If the former is required, generally some form of random-to-grid interpolation is required. Therefore all disciplines and applications requiring digital modelling involve an understanding of the relative accuracies of isarithm and random-to-grid interpolation. Additionally, while the surfaces and data samples utilised vary quite considerably, their characteristics must be carefully assessed in conjunction with the interpolation.

2.4 METHODS OF INTERPOLATION

From the previous discussion, it should be apparent that

> "The central problem of the DTM is one of interpolation" (Kraus, 1973, 2).

The interpolation of isarithms from data, like the generation of all the previously discussed applications in digital modelling, is invariably subdivided into a two-stage

procedure (Figure 2.1). In general, this involves the establishment of some intermediate grid through which isarithms are then threaded.

The first part - the establishment of some intermediate grid or triangulation net - is the aspect which represents the prime difference between the various methods of generating isarithms, and is the basic requirement within the previously discussed applications. Indeed Olea states that,

> "the kernel of any contouring scheme is the procedure used to generate the grid, since it consumes most of the computing time and the final results depend on its accuracy" (Olea, 1974, 695).

Because of this, random-to-grid interpolation methods will be examined first, and methods of triangulation generation will be considered subsequently.

The latter part - isarithm threading through the grid or triangulation net - also involves interpolation, but this is generally of the most basic form i.e. simple linear interpolation. This will be considered in the third sub-division of the section.

This section then describes the range of methods of isarithm interpolation generally available in digital modelling, as a preface to the following section which examines the programs, packages and subroutine libraries actually utilised in the research.

2.4.1 Random-to-grid Interpolation

In evaluating differing techniques of random-to-grid interpolation, several important aspects may be studied.

The range of the interpolating function at any instant forms the major difference between the differing techniques. 'Pointwise' methods involve the determination of different functional parameters for each and every point interpolated. A 'Patchwise' (2D) or 'Piecewise' (1D) approach involves the establishment of a series of local surfaces where several points are derived from each surface: however, if only one point is derived from the surface, then it is considered a pointwise procedure. A 'Global' range is based on a single function representing the whole surface from which the interpolated points are derived. Additionally, 'Hybrid' ranges may incorporate a blend of all three previous ranges, for example local functions may be derived which are 'summed' to give a more global function.

The interpolating function is the actual mathematical function that is fitted to each set of data to produce the interpolated point(s). This is generally some form of polynomial or trigonometric function, although other functions may be used.

Two aspects of the fitting procedure are of particular importance. First, the nature of the fit of the function may be varied. This may involve an exact fit or, alternatively, some form of approximate fit such as the Euclidean norm (least squares), the Gershgorin norm and the Chebyshev norm (minimax).

Finally, the method by which the original data points are weighted within the fitting procedure is important. Several forms of weighting and correlation/covariance functions are

frequently used.

2.4.1.1 Range Of The Interpolating Function.

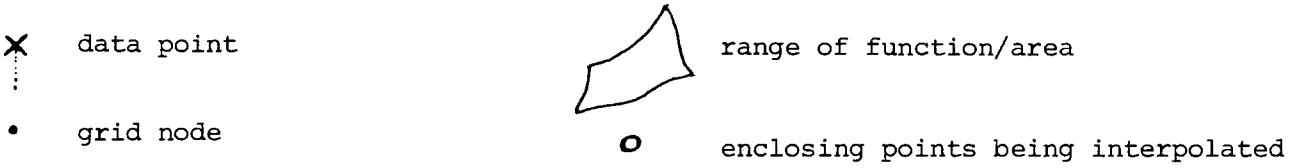
The range of the interpolating function varies between pointwise interpolation, patchwise/piecewise and global interpolation. Additionally, hybrid procedures provide an important variation. These are all summarised in Figure 2.9.

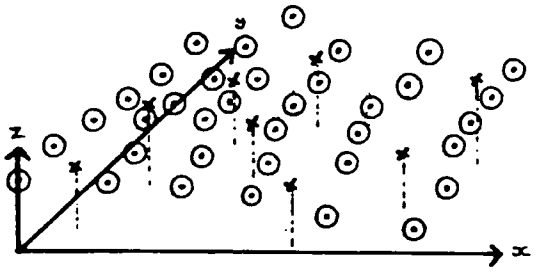
2.4.1.1.1 Pointwise Interpolation.

Pointwise interpolation involves the determination of separate functional parameters from selected neighbouring data points for each point to be interpolated (see Figure 2.9a). This method has many benefits.

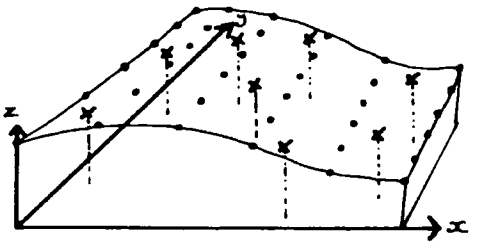
Generally, pointwise interpolation is used with some simple function, the simplest and commonest being either linear interpolation or a weighted average (eg. Connelly, 1971). This minimises the required computational effort and the necessary storage. Additionally, as each point is derived separately and sequentially, a continuous surface is generated with no obvious boundary problems except at the edges of two adjacent data areas.The interpolation is thus flexible, with the derivation of one point having no effect on the next. This has the benefit that where any data points may be subsequently added to the original data set, only a very limited number of local points must be re-interpolated.

The selection of those data points which participate in the interpolation of one point may become a time-consuming

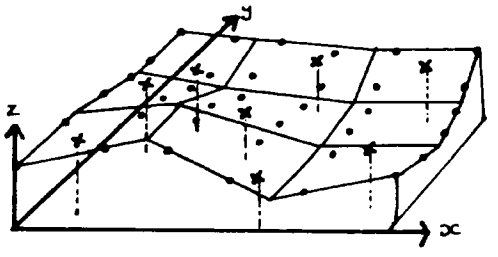




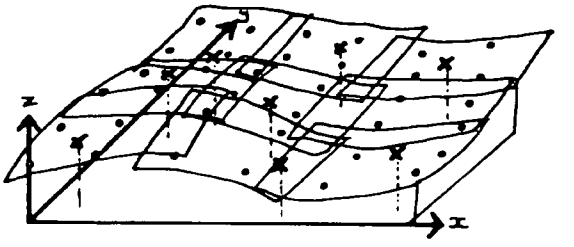
a) Pointwise



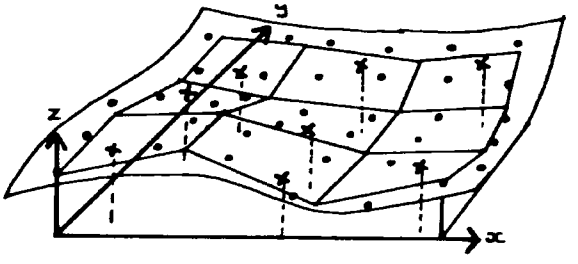
b) Global



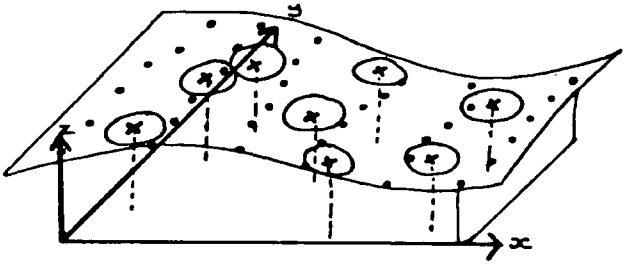
c) Patchwise - exact fit



d) Patchwise - overlapping fit



e) Hybrid - summation of surfaces



f) Hybrid - progressive fit

Figure 2.9 Graphical summary of the ranges of interpolation functions

process, though this may be greatly simplified if the data are systematically stored as they are input. Several searching techniques are utilised, generally the more sophisticated techniques being linked to the simpler interpolating algorithms and vice versa, based on the general belief that optimal interpolation may be obtained by using a complex function to overcome any limitations of the data distribution or by ensuring that the searched data for the interpolation are locally representative. Overall, there are three basic types of searching techniques, although combinations of these often occur. Figure 2.10 summarises the different techniques.

a. Area search. This is the simplest technique, resulting in all points within a fixed, generally unspecified area surrounding the interpolated point contributing to the interpolation (Worster, 1979 and University of Salford Computing Laboratory, 1979). While a circular search area is generally utilised to ensure an even symmetric search, other shapes may be used of which the square is the most common due to the simplicity of the required algorithm.

b. Number of nearest neighbours. A variation of the area search is to search for the 'n' nearest neighbours, where 'n' is usually user-specified. Generally, a normal default value for n is 8-10, although SYMAP V uses 7-14 (Shepard, 1968) and GPCP uses a default of 8 (Calcomp, 1973). The disadvantage of this technique, as with the area search, is that in selecting the nearest points, there is no control

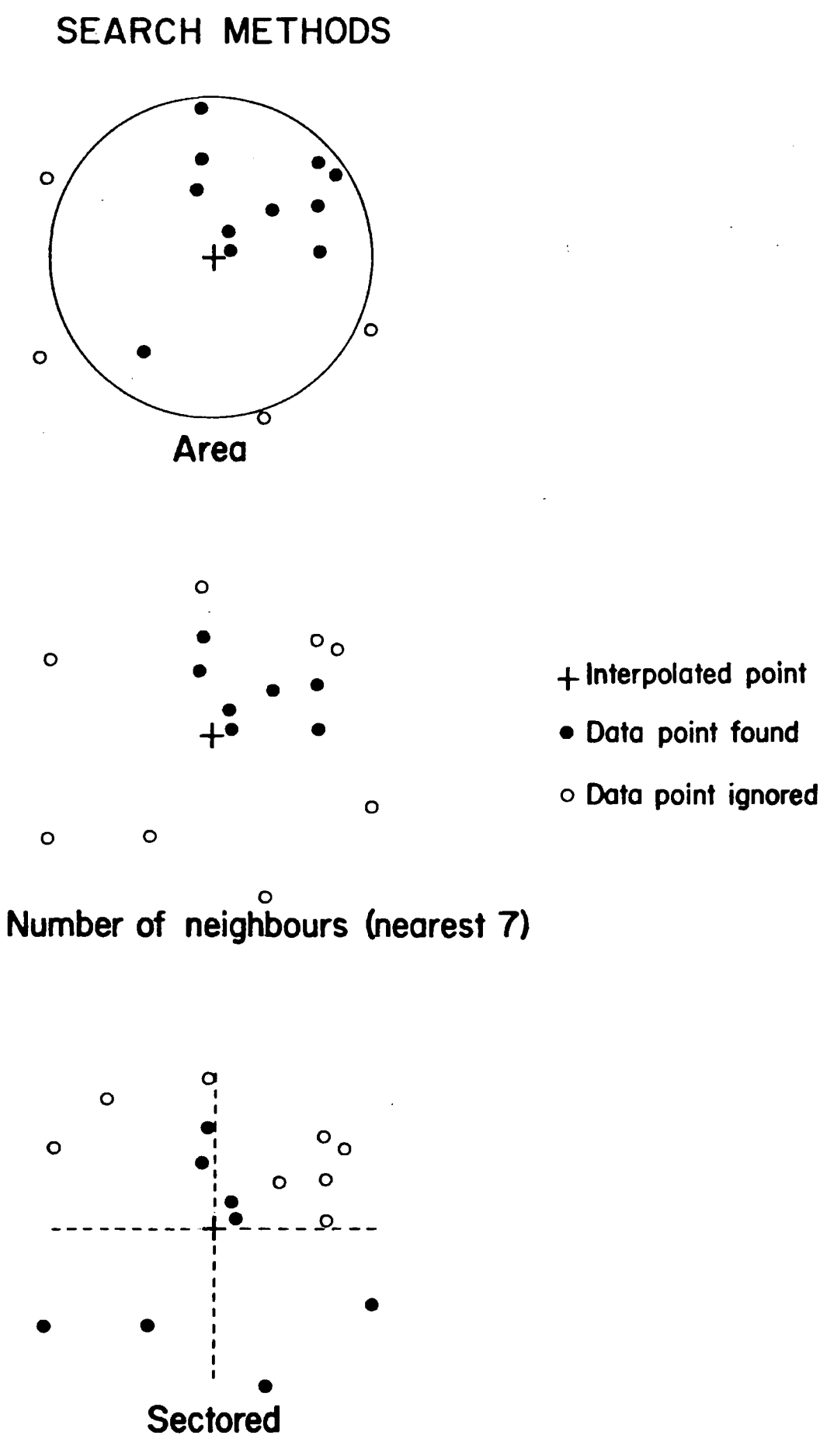


Figure 2.10 Search methods commonly used in interpolation methods

over their distribution.

c. Sectored nearest neighbours. This is a refinement of the number of nearest neighbours search where the area around the point is divided into sectors (e.g. SURFACE II GRAPHICS - Sampson, 1978 - gives the option of quadrant or octant search). The selection within each sector is performed as above, with generally some user-specified bounding distance limit, beyond which no search is made. This is most often required near the edges of the surface, where the search may be restricted to at the worst one quadrant. Generally, this search should result in a more even data point distribution than from b (above).

The problem of uneven data distribution influences the selection of participating data points and thus the interpolated value. Where a poor distribution of data exists, i.e. with obvious clusters and voids, problems may occur in the interpolation. This is the result of only considering a limited directional neighbourhood, and is manifested by the appearance of sudden spikes in the interpolated surface. Because the interpolated points are evaluated independently, the spikes are not smoothed out as perhaps in patchwise or (certainly) in global methods.

2.4.1.1.2 Global Interpolation.

Global interpolation involves the fitting of one function

to all the data points and the derivation of the interpolated values from this function (see Figure 2.9b). Where there is to be no data redundancy (i.e. the number of coefficients matches the number of data points), this may result (with large data sets) in problems of storage and computational time, as well as unwieldy oscillating functions. For this reason, and the fact that terrain has generally low autocorrelation, its application in DTMs is very limited and often results in a generalised, low-order, high data redundancy function being fitted (Howarth, 1969 and Norcliffe, 1969).

2.4.1.1.3 Patchwise Interpolation.

Patchwise techniques represent an attempt to combine the advantages of pointwise and global techniques. Essentially, the technique involves methodically dividing the total surface into several equi-sized and -shaped patches, and fitting separate functional parameters to each patch (see Figure 2.9c/d). The shape of the patches may be square, triangular, etc.. although in practice rectangles or squares are generally used. The size may allow for some overlap between patches, in which case there are common points within contiguous patches, or else result in 'exact fit' patches where each patch exactly abuts onto the next (Figure 2.9c). 'Exact fit' patches may result in sharp discontinuities in the resultant isarithms; this can be solved either by overlapping the patches (although at the expense of an increased computational effort - see Figure 2.9d) or by specifying continuity in the form of the function and considering each patch and its neighbours. Generally, the patches are larger than the areas of interest

for pointwise interpolation.

Patchwise interpolation is best considered as a continuum. If the whole area is covered by one patch, then global interpolation results; if only one point is solved per patch, then pointwise interpolation results; anywhere in between these two extremes is considered patchwise interpolation. Piecewise interpolation must be considered in conjunction with patchwise methods, since the technique is equivalent to patchwise interpolation confined to one dimension.

There are several advantages gained from using patchwise interpolation. Patchwise interpolation represents a compromise between the pointwise and global interpolation methods. Generally, low-order functions are used in patchwise methods, resulting in fewer unknowns being solved simultaneously than in global methods and, since fewer equations are established than with pointwise methods, fewer solutions are required. Once the unknowns are solved, the derived points may be computed more rapidly than in global methods. Thus, for the same degree of surface detail, patchwise interpolation is more efficient in terms of time and memory requirements than are global methods, and is more time-efficient than pointwise methods.

Use of the method also has disadvantages. It requires much more organisation of the data and of the processing than do pointwise or global methods. An efficient point storage scheme is still required and recomputation of several patches may be required if new points are added. Additionally, the subdivision of the surface into patches requires careful

organisation. In general, the function is sensitive to poor data distribution in patch corners. If poor point distribution occurs, then the problem of 'cracks' occurs at the boundary of the two patches (eg. Morrison, 1971).

Two inter-patch boundary situations occur. Patches may have matched boundaries, where cracks may be avoided by imposing some mathematical constraints on the function utilised, such as splining and enforcing general continuity of the first derivative. Alternatively, contiguous patches may have some common overlapping area; the size of any resultant crack then generally has an inverse correlation with the number of points and their distribution within such an overlap. Even with large overlaps, problems may occur and thus imposed constraints of some type are required. Braker (1975) has considered such constraints and classified them as follows: 'functionally explicit constraints' are where a large number of unknowns are used in the computation, 'functionally implicit constraints' are where the surface is represented by triangles and quadrilaterals and hence where disjoints are possible, and 'statistical constraints' are where residuals at the data points are kept to a minimum.

2.4.1.1.4 Hybrid Interpolation.

Many hybrid methods of interpolation have been derived in an attempt to combine the advantages of the different techniques. Some begin by defining localised areas and combining these to give a more global method (see Figure 2.9e); these come under the title, 'Summation of surfaces',

for example multiquadric analysis (Hardy, 1971, 1972a), Arthur's function of many variables (Arthur, 1965, 1973) and linear prediction (Schut, 1974). In these functions, a series of specialised low-order functions are fitted in a fashion similar to pointwise or patchwise techniques. Within multiquadric analysis, the precise form of the function may be based on an a priori knowledge of the data and surface. The interpolated points are derived from a series of functions which are weighted and summed, thus introducing some form of patchwise or global range.

Alternatively, as in Cole's Progressive Linear Fit (Cole, Jordan and Merriam, 1967 and Cole, 1969), a global fit may be first established and further processed at a local level (see Figure 2.9f). This generally utilises a low-order global polynomial and successive low-order local polynomials. However, none of these take into account any a priori knowledge of the surface, for example, overall orientation of the topography if lineated.

2.4.1.2 Interpolation Function.

Although there are many interpolation functions used in digital modelling, they are, when carefully examined, all derived from three basic classes. The major part of this discussion will concern the two most common techniques, those based on some form of mean or averaging process, and those which use some form of polynomial. The use of trigonometric series is mostly confined to global trend surfacing and macro-scale modelling.

2.4.1.2.1 Averaging.

> "The chief advantages of this technique ..(weighted average).. are that it always predicts reasonable heights and that the interpolated surface always passes through the control points" (Connelly, 1971, 59).

Averaging or arithmetic meaning is the most popular method used in pointwise interpolation. The formula for the process is

$$z = \left(\sum_{i=1}^{n} w_i z_i\right) / \left(\sum_{i=1}^{n} w_i\right),$$

where z = the value of the point being interpolated,

w_i = the weighting factor for each data point,

z_i = either the data point value or, more popularly, some derived estimate,

n = the selected sample size around the interpolated point.

The most basic version of the formula includes no weighting factor (i.e. a constant implicit weighting factor of 1) and represents a simple arithmetic mean of the surrounding data points. However, as nearer neighbours are generally more correlated in value than are distant neighbours, some weighting factor may be added (see 2.4.1.4). By definition, this function will only create values within the range of the data points locally averaged, and thus the average must be altered if values outside the local range of the data are to be generated. Several approaches exist to the incorporation of out-of-range values. In one of the more popular approaches, the z_i values used are derived from a first- or second-order polynomial surface fitted at each data point to their local data points, and solved at the interpolated point. These surfaces are provisionally derived before the grid

interpolation takes place, generally by a technique similar to the subsequent grid interpolation, and constrained so that they pass through the data point. This two-stage interpolation has the effect that since z_i may lie outside the range of the data points utilised, the mean may accordingly be outside the range of the local data points.

One of the most apparently complex yet inherently simple variations of the weighted averaging method is that of kriging. The term kriging has evolved in honour of D.G. Krige, a South African statistician and mining geologist who first used the method in the Klerksdorp gold field (Krige, 1966). Though it is impossible to assess kriging comprehensively in a few pages, a brief appraisal is given below.

Kriging is derived from the theory of regionalised variables and the concept that a regionalised variable is one that varies from place to place with apparent continuity, yet cannot be represented by a workable function (Matheron, 1971). Initially, the variance of the data is evaluated and co-variance along specifically orientated lines is established. From these, the semivariance and its graph, the variogram, are derived. Ordinarily, at least three variograms are constructed, two orthogonal ones and a third bisecting them. Kriging is then the process of estimating the value of a spatially distributed variable from adjacent values while considering these variograms which express the data inter-dependence.

UNIVERSAL KRIGING (Huijbregts and Matheron, 1970) is the most general form of operation, in which trends and

non-stationarity are present and the interpolation of points is estimated from irregularly distributed data. If it may be assumed that the data are stationary, the method may be simplified, and this is termed simply KRIGING. PUNCTUAL KRIGING (Olea, 1972 and 1974) assumes in addition that the data are isotropic i.e. that the three variograms are equivalent, and that the data are equi-spaced points on a grid. New points may be interpolated by a weighted average method, the weights being derived from functions fitted to the variograms (2.4.1.4). Manifestly, since a regular grid is required initially, this cannot be considered a form of random-to-grid interpolation!

2.4.1.2.2 Polynomials.

Polynomials are the basis for the most popular types of function in global and patchwise methods of interpolation. Davis (1973) claimed that their popularity is due primarily to the ease with which the function may be computed, while also due to the generally satisfactory fit that may be obtained even with lower-order functions. While the precise formula may be relatively complex, essentially it involves some assumption and the general formula:

$z = a_0$	0 order planar
$+a_1 x + a_2 y$	1 order linear
$+a_3 x^2 + a_4 y^2 + a_5 xy$	2 order quadratic
$+a_6 x^3 + a_7 y^3 + a_8 x^2 y + a_9 xy^2$	3 order cubic
$+a_{10} x^4 + a_{11} y^4 + a_{12} x^3 y + a_{13} x^2 y^2 + a_{14} xy^3$	4 order quartic
$+$etc.	

where z = the value of the point being derived,

$a_0, \cdots, a_{14}$ = the coefficients,

x,y = the x and y co-ordinates of the point.

The number of coefficients may be calculated from the general formula,

n = ((order + 1).(order + 2)/2).

As a general 'rule-of-thumb', pointwise interpolation involves orders 0 and 1, patchwise and hybrid interpolation orders 1,2 and 3, and global interpolation from fourth- to tenth-orders.

There are several specialised polynomial techniques which are worthy of being examined more fully.

2.4.1.2.2.1 Multiquadric analysis.

A concise definition of multiquadric analysis is given by Hardy

> "Multiquadric analysis is a method of deriving equations of irregular surfaces from a summation of regular surfaces, each associated with a particular co-ordinated control point in the irregular surface" (Hardy, 1972b, 398).

The theory of multiquadric analysis is based on the multiple derivation of quadrics. A quadric is a locus in three-dimensional space defined by a quadratic equation in x, y, and z, i.e. an equation of the form

$$a_0 x^2 + a_1 y^2 + a_2 z^2 + 2a_3 yz + 2a_4 zx + 2a_5 xy + 2a_6 x + 2a_7 y + 2a_8 z + a_9 = 0,$$

where x,y,z = the three-dimensional co-ordinates

of a point,

$a_0, \ldots, a_q$ = the constants.

Essentially, the possible resulting surfaces may be termed degenerate or non-degenerate. The generate cases include the cone and cylinder. Alternatively, there are five non-degenerate quadrics: the ellipsoid; hyperboloid of one sheet and two sheets; hyperbolic paraboloid and the elliptic or circular paraboloid. Various workers (Brooks, 1970; Hardy, 1971 and 1972a and Shaw and Lynn, 1972) have popularised this approach and, by using a hybrid approach involving summations of local quadrics centred on each data point, they have obtained very promising results.

In the general formula, Hardy (1975) represents his multiquadric surface by

$$\sum_{i=1}^{n} a_i (q(x_i, y_i, x, y,)) = z,$$

where z = a function of x and y resulting from the summation of a single class of quadric surfaces q,

x_i, y_i = the location of the vertical axis of symmetry of each quadric term,

n = the number of points,

a_i = the coefficient which determines the algebraic sign and flatness of the quadric term.

From this, the following quadrics are commonly derived.

1) $\sum_{i=1}^{n} a_i((x_i-x)^2+(y_i-y)^2+a)^{1/2} = z$,

A series of circular hyperboloids;

2) $\sum_{i=1}^{n} a_i((x_i-x)^2+(y_i-y)^2+a) = z$,

A series of circular paraboloids;

3) $\sum_{i=1}^{n} a_i((x_i-x)^2+a)^{1/2} = z$,

A series of hyperboloid sections;

4) $\sum_{i=1}^{n} a_i((x_i-x)^2+(y_i-y)^2)^{1/2} = z$,

A series of cones;

where a = an arbitrary constant.

With this method, there are two main disadvantages. As a quadric is centered on each of the data points, it is clear that, to represent the surface correctly, the data should always be located on the peaks or depressions of the surfaces - as is desirable when using the averaging process. In addition, by the very nature of the computation, a matrix must be inverted whose dimensions are equivalent to the number of data points. Naturally problems of computer storage and time occur as the number of data points increases, and this eventually restricts the size of the input data set. This method has therefore low time- and storage-efficiency.

2.4.1.2.2.2 Splines.

The mathematical spline is directly analogous to the device used by draughtsmen to draw a smooth curve (Ahlberg, Nilson and Walsh, 1967 and Greville, 1967). This consists of a strip or rod of some flexible material, to which weights are

attached so that it can be constrained to pass through or near certain plotting points. It has been shown that the curve produced by a spline is a cubic spline function (Schoenberg, 1946).

The spline is a piecewise polynomial function satisfying certain conditions regarding continuity of the function and its derivatives. In a strictly increasing sequence of real numbers $x_1, x_2, \ldots, x_n$ a spline function S(x) of degree m with knots $x_1, x_2, \ldots x_n$ is a function defined on the entire real line having the properties that,

- in each interval (x_i, x_{i+1}) for $i = 0, 1, \ldots, n$, S(x) is given by some polynomial of degree m or less,
- S(x) and its derivative of orders 1, 2, ..., m-1 are continuous everywhere.

While, in general, S(x) is given by different polynomials in adjoining intervals, the definition does not require this and may be given by a single polynomial on the entire real line. Thus the general one-dimensional form of the spline S(x) of degree m with the knots $x_1, x_2, \ldots, x_n$ is

$$S(x) = p(x) + \sum_{i=1}^{n} a_i (x - x_i)_+^m ,$$

where $p \in \pi_m$

π_m denotes the class of polynomials of degree m or less, (Greville, 1969).

As with the basic polynomial the order and dimensionality are infinitely flexible, their choice being linked to the application. In isarithm interpolation, when the actual isarithm has to be smoothed, frequently a one-dimensional piecewise cubic spline is utilised (McConalogue, 1970 and

Cline, 1974a and 1974b). In surface interpolation, bi-dimensional patchwise quadratic splines are often used although, more frequently, the cubic spline is selected (Hayes and Halliday, 1972). Very rarely in surface modelling are fourth- and higher-order splines used, but this is inherently possible in the function range. Bi-dimensional splines are only used in patchwise interpolation when continuity across patches is required (eg. Holroyd and Bhattacharyya, 1970 and Davis and David, 1980).

There are further mutations of the natural spline. Hayes and Halliday (1972) consider the 'B-spline' (or Fundamental spline), which has non-zero gradient only over four adjacent intervals between knots, to be superior for surface modelling. Hartley and Judd (1980) derive 'Bezier' splines, an extension of bi-dimensional B-splines, and consider them superior for engineering design modelling. In general and irrespective of the type of spline, the splines are a piece- or patchwise method allowing smoothness of surface or line fit to localised data.

2.4.1.2.2.3 Orthogonal polynomials.

Orthogonal polynomials are frequently utilised in preference to non-orthogonal polynomials. When normal non-orthogonal polynomials are used two problems often occur. The first is that normal equations in the approximation frequently become ill-conditioned causing problems in their solution which lead to errors in the results. Edwards claims

that

> "problems of ill-conditioning can be overcome by the use of orthogonal polynomials (Edwards, 1972, 586).

More importantly, it causes serious problems in computer processing, as by increasing the number of points and (especially) the order of the polynomial, there is a considerable increase in the computing time and storage required. The use of orthogonal polynomials solves both of these difficulties.

Orthogonal polynomials may be obtained in many ways, although the Gram-Schmidt orthogonalisation procedure is frequently used (Conte and de Boor, 1972 and Forsythe, 1957).

Let (a,b) be a given interval, and let w(x) be a given function defined (at least) on (a,b) and positive there. The scalar product $\langle g,h \rangle$ of any two functions g(x) and h(x) (defined on (a,b)) may be defined as,

$$\langle g,h \rangle = \int_a^b g(x).h(x).w(x)dx$$

$$\langle g,h \rangle = \sum_{i=1}^{n} g(x_i).h(x_i).w(x_i).$$

With the scalar product of two functions defined it may be said that the two functions g(x) and h(x) are orthogonal (to each other) in the case

$$\langle g,h \rangle = 0.$$

Further, it may be said that $P_0(x)$, $P_1(x)$, $P_2(x)$,... is a (finite or infinite) sequence of orthogonal polynomials provided that $P_i(x)$ are all orthogonal to each other and each $P_i(x)$ is a polynomial of exact degree i (Conte and de Boor, 1972).

2.4.1.2.2.4 Fourier series.

Fourier series are probably the most widely used alternative function to polynomials and averaging, the Double Fourier function being analogous to the bi-dimensional polynomial.

In the single dimension, the (single) Fourier series is represented by,

$$z_i = \sum_{k=1}^{\infty} (a_k \cdot \cos(\frac{2k\pi x_i}{\lambda}) + b_k \cdot \sin(\frac{2k\pi x_i}{\lambda}),$$

where z_i = the amplitude at a point x_i,

λ = the fundamental wavelength,

a_k, b_k = constants,

π = pi, the value 3.1416,

k_{max} = the number of harmonics in the x-direction.

This is easily converted into the bi-dimensional Double Fourier series,

$$z_{ij} = \sum_{k=1}^{\infty}\sum_{l=1}^{\infty} a_{kl} \cdot \cos(\frac{2k\pi x_i}{\lambda_1}) \cdot \cos(\frac{2l\pi y_j}{\lambda_2})$$

$$+ \sum_{k=1}^{\infty}\sum_{l=1}^{\infty} b_{kl} \cdot \cos(\frac{2k\pi x_i}{\lambda_1}) \cdot \sin(\frac{2l\pi y_j}{\lambda_2})$$

$$+ \sum_{k=1}^{\infty}\sum_{l=1}^{\infty} c_{kl} \cdot \sin(\frac{2k\pi x_i}{\lambda_1}) \cdot \cos(\frac{2l\pi y_j}{\lambda_2})$$

$$+ \sum_{k=1}^{\infty}\sum_{l=1}^{\infty} d_{kl} \cdot \sin(\frac{2k\pi x_i}{\lambda_1}) \cdot \sin(\frac{2l\pi y_j}{\lambda_2}),$$

where z_{ij} = the amplitude at a point x_i, y_j

$a_{kl}, b_{kl}, c_{kl}, d_{kl}$ = constants

λ_1, λ_2 = the fundamental wavelengths in the x and y directions,

k_{max}, l_{max} = the number of harmonics in the x y directions.

This in turn may be simplified by setting,

$$C_k = \cos\left(\frac{2k\pi x_i}{\lambda_1}\right)$$

$$C_l = \cos\left(\frac{2l\pi y_j}{\lambda_2}\right)$$

$$S_k = \sin\left(\frac{2k\pi x_i}{\lambda_1}\right)$$

$$S_l = \sin\left(\frac{2l\pi y_j}{\lambda_2}\right)$$

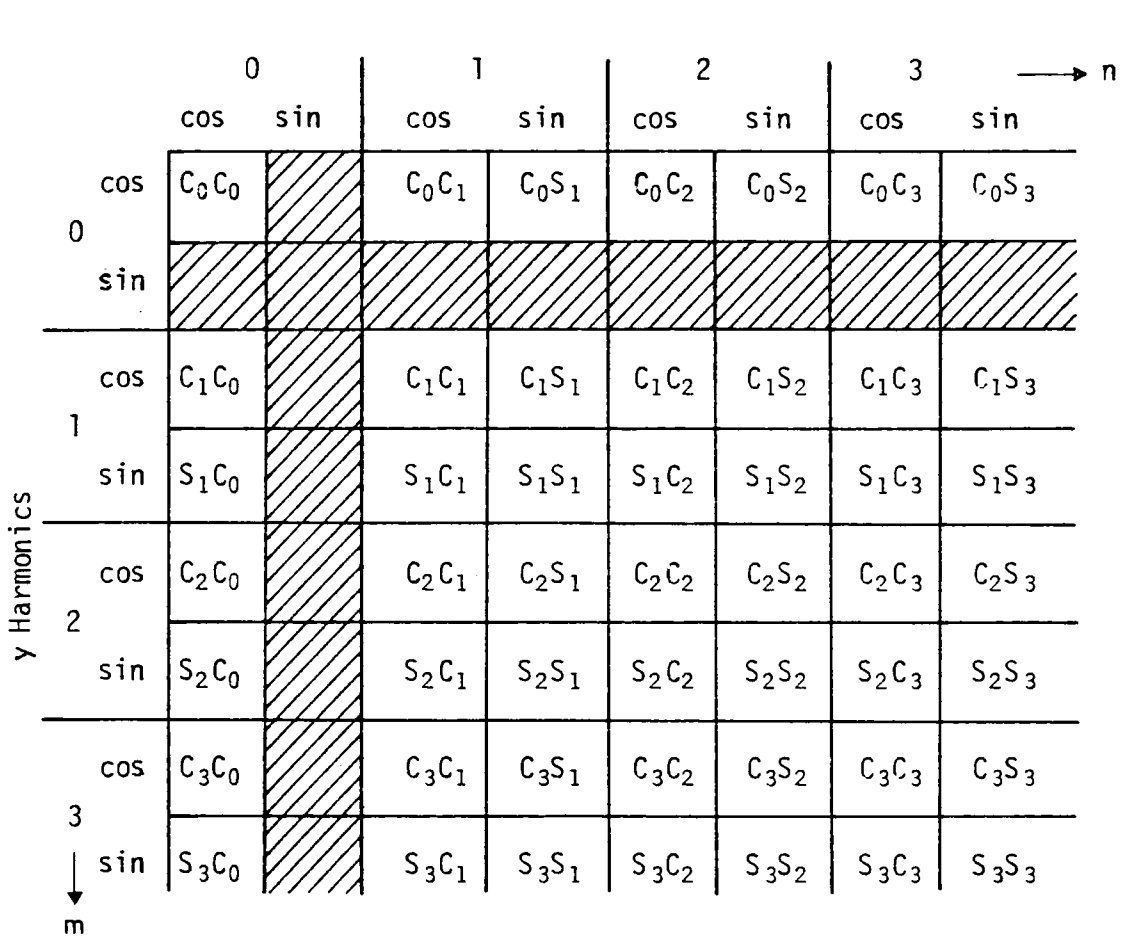
and rewritten as,

$$z_{ij} = \sum_{k=1}^{\infty}\sum_{l=1}^{\infty}\left(a_{kl}C_kC_l + b_{kl}C_kS_l + c_{kl}S_kC_l + d_{kl}S_kS_l\right).$$

This is expressed more objectively in Figure 2.11.

When the harmonic of a particular term is zero, the sine is zero and the cross-product vanishes (see the shaded area in Figure 2.11). Thus, assuming equal harmonics in both co-ordinate axes directions, block 0 (zero harmonics in both directions) requires 1 coefficient to be solved; block 1 (one harmonic in each direction) requires 8 additional coefficients; block 2 (two harmonics in each direction) requires 16 additional coefficients, etc. The number of

Figure 2.11 Double Fourier series coefficients arranged according to wavelength. Shaded coefficients are zero. (after James, 1966)



Order/Harmonics	Number of coefficients	
	Polynomial	Fourier
0	1	1
1	3	9
2	6	25
3	10	49
4	15	81
5	21	121

Table 2.1 Comparison of number of coefficients in polynomial and Fourier series.

coefficients therefore increases dramatically faster than for increasing polynomial order (Table 2.1). This represents a major disadvantage in the solution of Fourier series, imposing severe demands on computer memory and in round-off errors which become increasingly troublesome in the inversion of the necessarily large matrices. The development of the Fast Fourier Transform (FFT), an algorithm resulting in a manifold reduction of computational time and making the analysis of large data sets practical (Barrett and Rhind, 1974), has therefore been of considerable significance.

Davis states that Fourier series are,

> "more appropriate approximating functions than power series polynomials if the data seem to contain spatially repetitive elements" (Davis, 1973, 364).

Some natural surfaces have periodicity, but often this is restricted to a few situations at a limited number of scales (eg. drumlins at small scale, fold mountains at a large scale and outside the study of topography, the Earth's magnetic field in macro-scale modelling). The majority of natural surfaces have no obvious periodicity, and thus it is felt that Fourier series have little place in this thesis as regards interpolation and have therefore not been examined further. However, their use is discussed further in the context of spectral analysis (see 4.3.4).

2.4.1.3 Fit.

In order to fit any function $f(x,y)$ to a sample of data points the number of coefficients must not be greater than the number of data points. If the number of points matches the

number of coefficients, then the fit is said to be exact. If the number of points is greater than the number of coefficients, then the fit is said to be approximate, in which situation there are three norms in most frequent use. The fits may be summarised as:

2.4.1.3.1 Exact Fit.

Exact fits are most popular in pointwise and generally low-order interpolating functions, where the value of the function at the data point is equivalent to the observed value, i.e.

$$E_i = f(x_i, y_i) - z_i$$
$$= 0;$$

where E_i = the error at point i,

z_i = the observation at point i ,

$f(x_i, y_i)$ = the computed value at point i.

2.4.1.3.2 Approximate Fit.

Approximate fits are most popular in global and patchwise interpolation where data redundancy occurs or is assumed to occur, and thus the function f(x,y) assumes values at the data points approximately equal to the observed values. The error or residual at every data point may again be derived from

$$E_i = f(x_i, y_i) - z_i .$$

In order to obtain a good approximation, the errors must be kept within certain bounds by some error criterion (Berztiss, 1964), and in general terms the modulus of the error |E|

should be kept to a minimum. There are three norms used to satisfy this condition, in each case the best estimate is one satisfying the requirement, $||E|| = \min$,

2.4.1.3.2.1 The Gershgorin norm satisfies the requirement

$$||E||_1 = \sum_{i=1}^{n} |E_i|$$

and is generally little used in digital modelling. Indeed Berztiss

> "is unable to think of anything to recommend the Gershgorin norm" (Berztiss 1964, 204)

for the general problem considered here.

2.4.1.3.2.2 The Chebyshev norm or 'minimax' condition

satisfies the requirement

$$||E||_\infty = \max|E_i|, \quad i=1,\ldots,n .$$

This reduces the maximum deviation to a minimum at the cost of an increased average square error, and thus is more suitable than the other approximate fits where the observations are of a greater precision. However, it is often suggested that, the determination of f(x,y) according to the 'minimax' criterion is rather complicated even in one dimension (Conte and de Boor, 1972).

2.4.1.3.2.3 The Euclidean norm or 'least squares' condition

satisfies the requirement

$$||E||_2 = \left(\sum_{i=1}^{n} E_i^2\right)^{1/2} ,$$

permitting larger deviations at some points, but keeping

square error to a minimum. Least squares is simpler to compute and more commonly used than is the Chebyshev norm. Usually little is known about the distribution of errors in data and the assumption is generally made that the errors are normally distributed. If this assumption is valid, then least squares estimates must be considered optimum for most applications. Since the other methods of approximate fit are seldom (if ever) used, further discussion will be curtailed.

2.4.1.4 Weighting/Covariance Function.

Having considered the function, its range, and method of fit, the method of taking into account the relative importance of the data points must be examined. 'Relative importance' is a complex factor to evaluate and thus is generally established a priori although increasingly a posteriori evaluation is being used, especially in hybrid interpolation. The net result of either evaluation is the establishment of a weighting function which is then applied to the data directly during interpolation and/or to intermediate functions in most hybrid forms of interpolation.

The use of a weighting function to give more weight to certain observations than another is a simple enough concept, for example most pointwise interpolation methods use some well-used distance decay function. Howevef, the more recent patchwise and hybrid methods, especially those involving summation of surfaces, and other methods involving the theory of stationary random variables (for example kriging) use data-generated functions or data covariance functions which,

if normalised, are termed correlation functions. These data-derived functions are usually fitted to some general mathematical expression, involving exponential trigonometric or distance elements, and a final weighting function established. Thus a unique weight may be established for a certain data point from the mathematical expression. In general, the pre-conditions of such a function are that,

a. it must be positive at d=0,

b. it must have its maximum at d=0,

c. it must converge to zero for $d^{\rightarrow\infty}$,

d. it usually will be positive and continuous,

(where d is the distance between the point to be interpolated and the data point).

Thus, in summary, either the weight at a point is established from a known function, or the weight is derived from a general function fitted to the covariance of the data. Both methods utilise the same types of functions, but the actual form is established <u>a priori</u> in the former and <u>a posteriori</u> in the latter.

The number of possible weighting functions used during interpolation is infinite. Very broadly, they may be

categorised into those involving distance only, those involving an exponential function and those involving an additional trigonometric function. It should be stressed however, that this grouping does not involve discrete classes.

2.4.1.4.1 Distance Functions.

The simplest method of weighting involves the equation of basic form,

$w = 1/d^n$,

where w = the weight of the data point,

d = the horizontal distance between the data point and the point to be interpolated,

n = the power (usually of range 0.5 to 4) .

Crain and Bhattacharyya (1967) and Crain (1970) use the function with order n=0.5, Turner and Miles (1968) with n=1, McIntyre, Pollard and Smith (1968) with n=2 and Pelto, Elkins and Boyd (1968) claim that n=4 "has proved successful". Essentially, the effect of decreasing n in this and most other functions is to increase smoothing. Thus $1/d^4$ has a faster decay of weight with increasing distance than 1/d (see Figure 2.12a).

Alternatively, many workers have derived ratio functions in a similar format to the distance functions (Figure 2.12b). Thus,

$w = 1/r^n$

where r = the ratio (d/the maximum distance to be considered),

with n=2,4,6, Hoeli program (Schut, 1976),

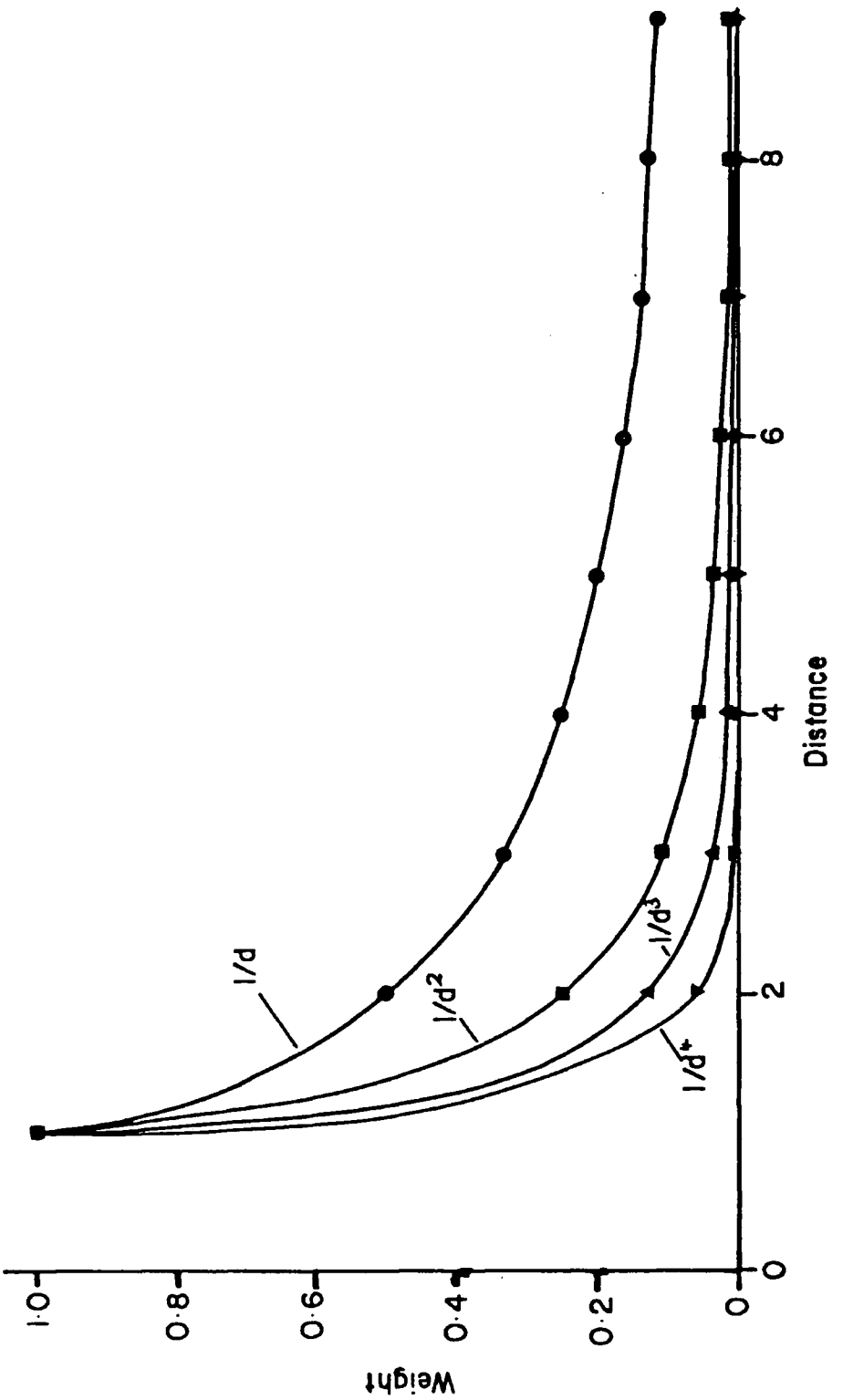
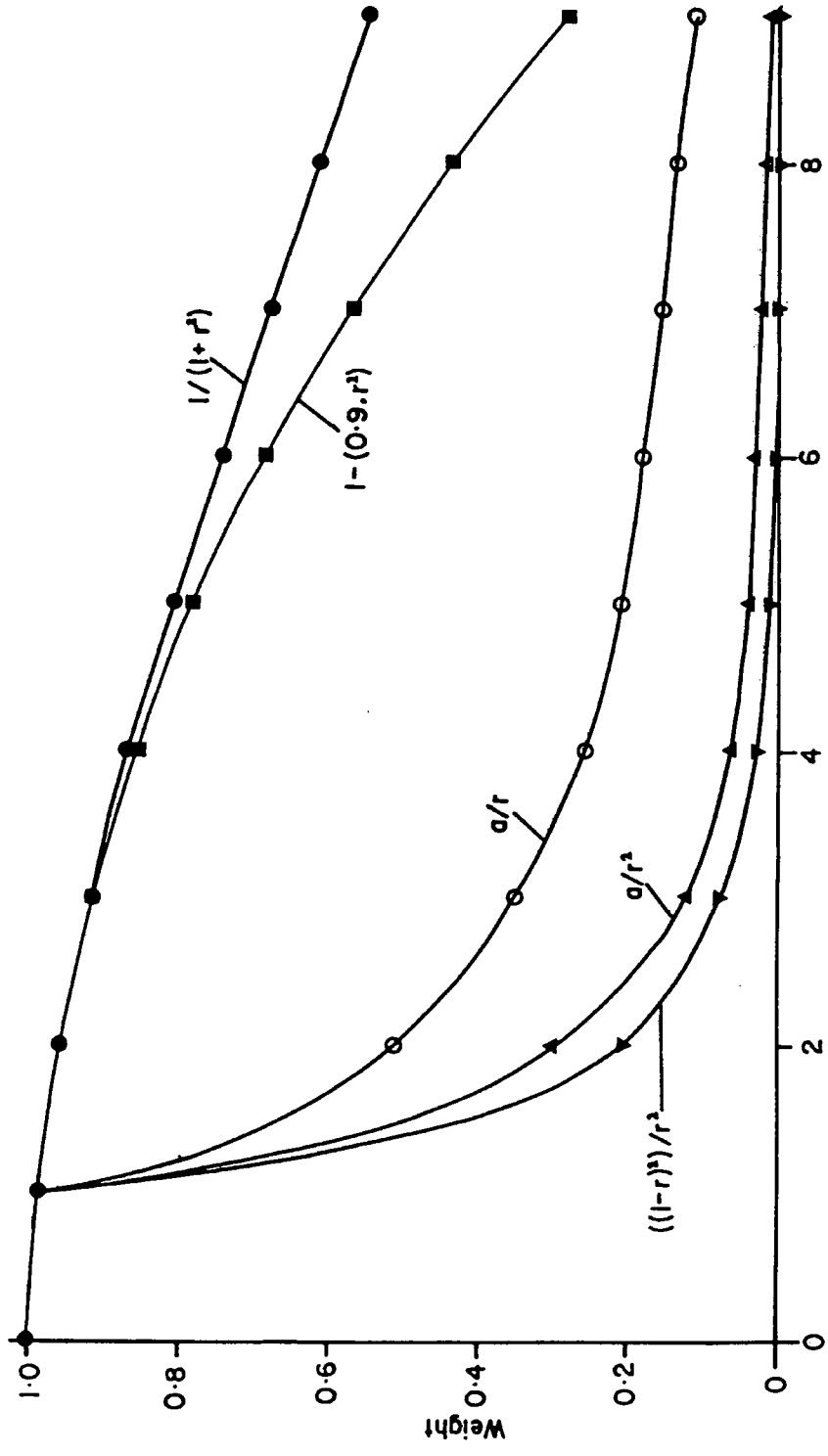


Figure 2.12a Decay with distance of distance - based weighting functions.



Fibure 2.12b Decay with distance of ratio-based weighting functions

n=2 , (for second-order surfaces),

n=4 , (for first-order surfaces),

IBM Germany (Schut, 1976),

n=1,2,3 , IBM Sweden,

n=3 , Swedish road service.

This basic function produces a similar weight decay to the distance only functions (Figure 2.12a). However the decay of the function may be modified by altering its algebraic complexity. Connelly's ECU package used a function,

$$w = (1-r)^2 / r^2 .$$

This shows a faster decay than the simpler ratio functions, (Figure 2.12b). However, increased smoothing was generated by applying the condition that r=0.3 if r<0.3.

Similarly Schut (1976) describes other, more complex functions:

$w = (1-r)^3 . (1-r^2)^5 / r^n$; n=1,2,

(this is reported to produce good results without excessive smoothing),

$w = 2 (1-r)^2$; $r \geqslant 0.5$,

$w = 1-2r^2$; $r \leqslant 0.5$,

(it is suggested that these are both suitable functions for very strong smoothing),

$w = 1-(0.9r^2)$, (Figure 2.12b),

$w = 1/(1+r^n)$; n=1,2, (Figure 2.12b),

(these both produce much heavier smoothing than the simpler ratio functions).

Other workers mention,

$w = (d^2+a)^p$; a = a constant, p = 0.5,1 (Hardy, 1971),

$w = 12/(d+1)$,

$w = 24/(d^2+1)$, (both Newton, 1968),

$w = 1/(1+d^n)$; $n>0$ (Braker, 1975),

$w = 1-r^2$, (Arthur, 1965).

Distance and ratio-based functions are by far the most common in use. The above discussion is not meant to be exhaustive, but rather to stress that there is such a range and that individual weighting functions may produce markedly similar or differing results. In applying such an approach, each function must be examined in conjunction with the degree of smoothing and weight decay that is required or is appropriate to a particular data set and application.

2.4.1.4.2 Exponential Function.

There are also many examples of the use of exponential functions, where benefits have been described as easing possible problems of covariance matrix singularity. The basic form of the functions is similar to the distance and ratio functions, although the variations utilised in practice are not so numerous.

The basic form is,

$w = \exp(-a.d)$; where $a>0$, (Yaglom, 1965).

This produces a gentler and smoother decay than with the basic distance and ratio functions, especially where a tends to 0, (Figure 2.12c).

This has been further developed within the SCOP package to

$w = \exp(-a.d^2)$; where $a>0$, (Schut, 1976).

This produces a more rapid decay, for comparable values of a,

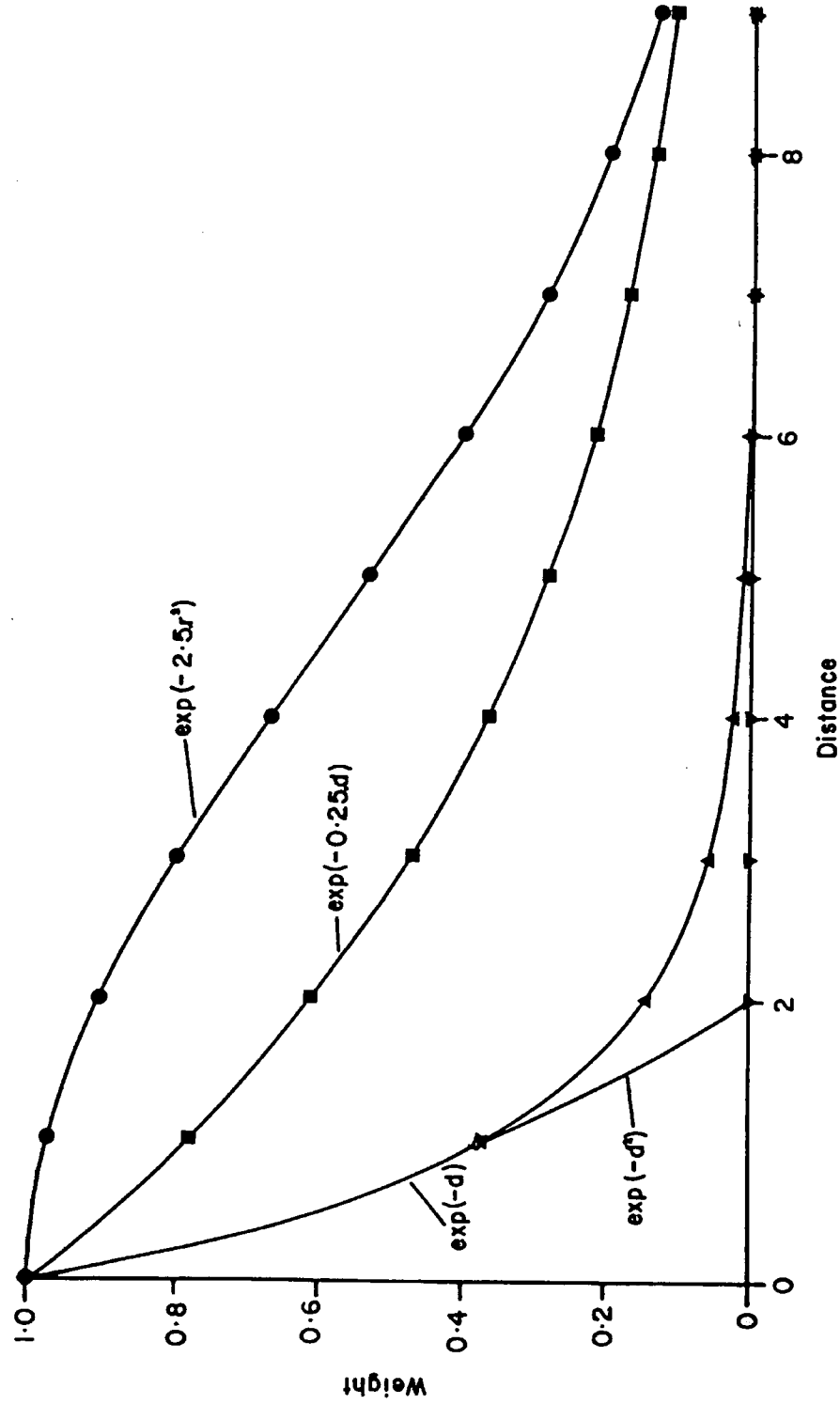


Figure 2.12c Decay with distance of exponential - based weighting functions

than the basic form and thus would be more suitable to less correlated data.

Simlarly, distance ratios have evolved, e.g.,

$w = \exp(-a.r^2)$;

where a = constant = 14 or 20 (Schut, 1976),

or = 2.5 (Arthur, 1973).

In the latter version, which produces a very gentle decay (Figure 2.12c), Arthur also derives the ratio from,

r = d / (average distance between reference points).

Other variations include those of Lauer,

$w = \exp(-a.r^b)$;

where a > 0,

b (he claims) to be optimum at 1.2, (Schut, 1976), and Braker (1975),

$w = \exp(-a.u^2)$;

where $u = r.(b+(1-b).r)$,

b > 0.

2.4.1.4.3 Trigonometric Functions.

Finally, some workers have gone a stage further and included trigonometric terms in the weighting functions, for example,

$w = \exp(-a.d).\cos(b.d)$; where a>0, b>0, (Yaglom, 1965),

$w = \exp(-a.d^2).\cos(b.d)$; where a>0, b>0, (Schut, 1974).

Examples of these two functions are given in Figure 2.12d. An initial aspect of these two functions is that they have a very rapid decay close to d=0, i.e. over short distances. More importantly, the introduction of the trigonometric term has

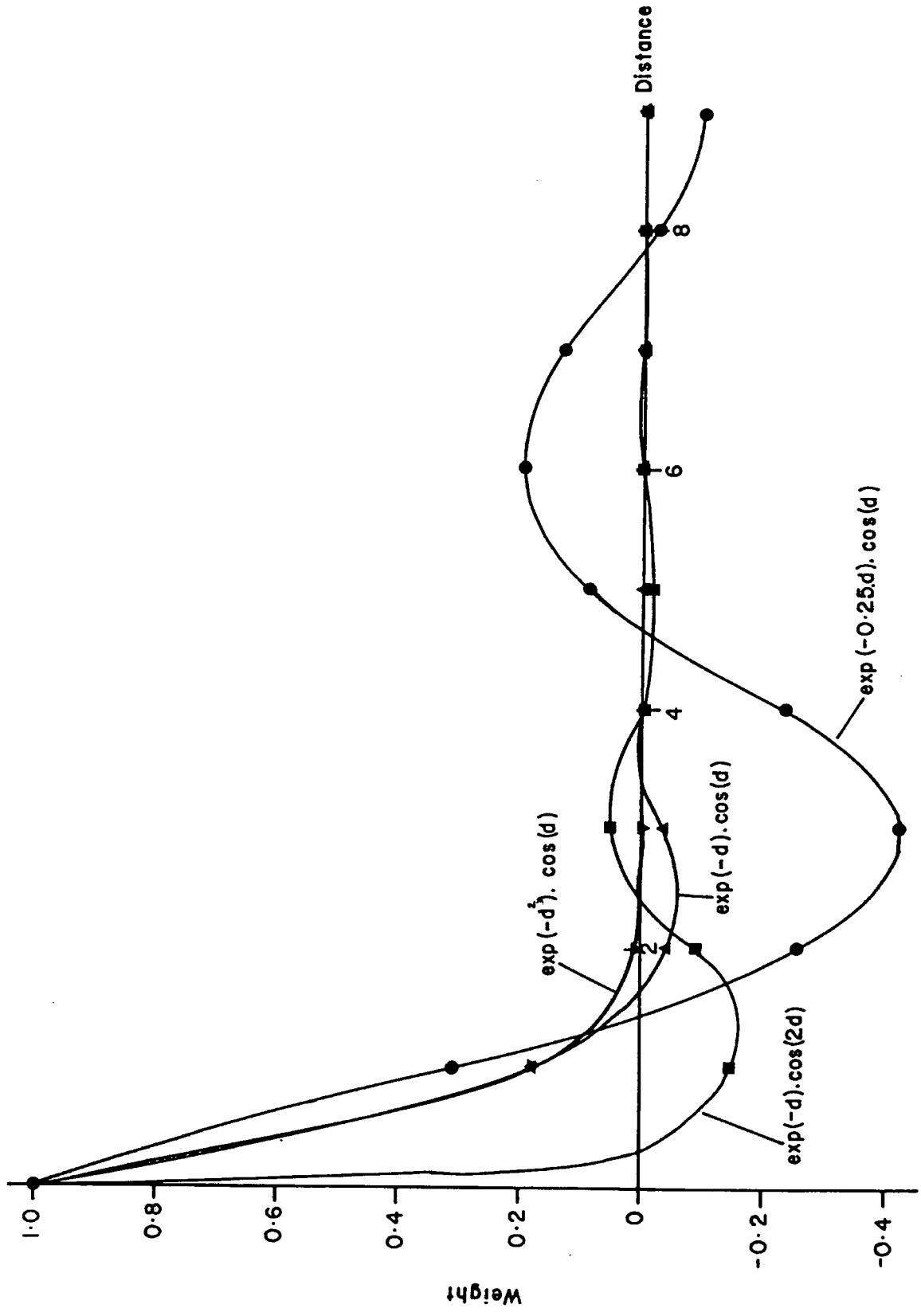


Figure 2.12d Decay with distance of trigonometric - based weighting functions

introduced negative weighting. This may be quite permissible given a periodic surface of peaks and troughs, where a trough has a negative correlation with a peak; its use does, however, presuppose considerable a priori knowledge of the surface.

In summary, a wide range of mathematical functions exist and have been used for weighting data within interpolation. Their application results in distinctive variations of smoothing and weighting and hence requires careful consideration by the user - whether or not he has any control over the function used.

2.4.2 Triangulation Generation

A triangulation mesh represents the most robust yet the simplest solution for isarithm interpolation. Where no interpolation has taken place in the process of its generation, it has the advantage that it honours every data point, and should therefore, where the observations have been sufficiently consistent, produce isarithms close to reality. Essentially there are two starting points. The input to the triangulation may be a grid (irregular, regular, etc.), or a random scatter of points.

2.4.2.1 Grid Triangulation.

The rationale for producing a grid-based triangulation will be stressed in a later section. There are two methods by which a grid may be modified to a triangular mesh.

2.4.2.1.1 Diagonal Method.

The diagonal method (see 'GHOST' - A1.3.4) involves dividing each grid cell into two triangles using one of the cell diagonals (Figure 2.13a). This method has the advantage of ease of programming, data manipulation and storage, and additionally does not involve interpolation with its resultant loss of accuracy. Depending on which diagonal is chosen, the shape of the contour(s) may change (Figure 2.13a).

2.4.2.1.2 Centre-point Method.

The centre-point method (see 'GINOSURF' - A1.3.5 and 'CONSYS' - A1.3.1) involves some form of secondary interpolation, generally an averaging or weighted averaging process. In any grid cell, the centre point is established by either an average or a weighted average of the four surrounding nodes. Thus a value within the range of the four nodes will be established, which is then used with the four nodes to generate four triangles (Figure 2.13b). The advantage of this method is that it makes the triangulation resolution even finer than the diagonal method, which should result in shorter and therefore visibly smoother isarithms. However, this is at the expense of a possible decrease in accuracy due to secondary interpolation.

2.4.2.2 Random Data Triangulation.

Random data triangulation is one of the most popular methods of generating isarithms from random data as it honours every data point if direct linear interpolation of contours is

Figure 2.13 **The effect of various grid-to-triangulation methods**

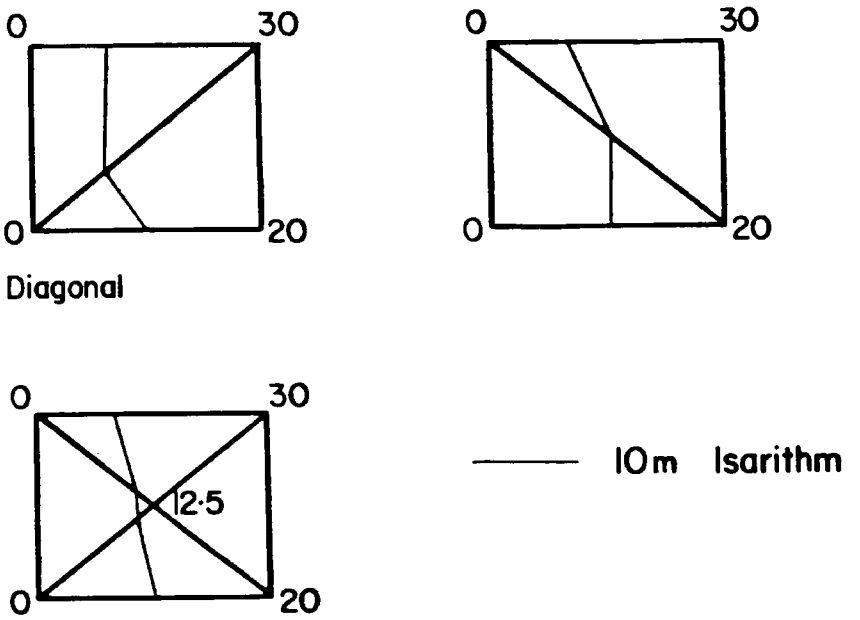
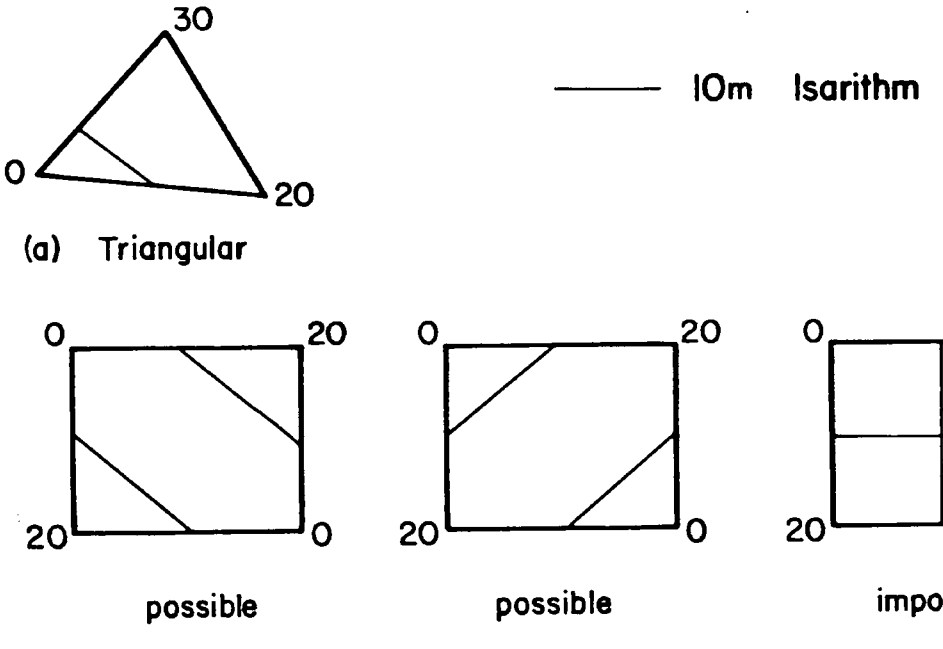


Figure 2.14 **Contour threading ambiguity**



subsequently used. In relation, especially to topographic structures, it additionally allows a simple method of including breaklines and thus controlling the definition of the irregularities in the surface. However,

> "any triangulation that is to be used for the basis of isarithm map production must have the properties of stability, equilateralness, and non-intersection" (McCullagh and Ross, 1980, 93).

The Delaunay triangulation meets these needs and is therefore most frequently used. The logistics of its generation however may be quite varied, indeed there are four methods generally used for its derivation (Elfick, 1979, 24).

a. IBM published a method which first derived a series of regular polygons which were later broken down into triangles.

b. A generally used method is to build on an advancing front, for example from left to right, until all the data points are used.

c. Yoeli (1977) generates a triangular mesh by constructing triangles progressively from the shortest unused line segments, while allowing for breaklines. This method is also employed within the Green and Sibson (1978) routines called by 'DTC' in this research.

d. A fourth method, used by Elfick (1979) and within

MOSS, builds the triangular mesh progressively on the sides of existing triangles.

2.4.3 Isarithm Interpolation

> "It is important that the contouring package recognises and honours the surface being contoured" (Monahan, 1971, 350).

There are two forms of initial input to the isarithm interpolation stage. A common input in civil engineering is a random triangulation - as previously discussed. Alternatively, the input is equally commonly a rectangular grid. The actual form of the rectangular grid is trivial as the logic of interpolation is the same. Indeed, excluding the frequent derivation of a triangulation net as an intermediate stage in interpolating isarithms from a grid, the two forms of input to the isarithm process are, in broad terms, similarly treated, although the logistics are somewhat different. Essentially, the differences (excluding input) revolve around attempts to ensure continuity along the isarithms.

2.4.3.1 Isarithm Interpolation By Triangulation.

Assuming a triangulation network of the data points has been established, two options exist for the isarithm threading.

2.4.3.1.1 Direct Isarithm Threading.

Direct isarithm threading is the most popular form of isarithm generation and is based on linear interpolation. For each isarithm level, each triangle is sequentially searched until one is found whose 'z' range encompasses the isarithm level under interest. Initially the bounding triangles are searched; once all the isarithms which start at an edge are established, the internal triangles are searched.

Once the starting edge has been established, a simple linear interpolation is performed to derive the co-ordinates of the isarithm's starting point. The point of exit is thereafter derived by first establishing the side of exit and then, by linear interpolation, establishing the co-ordinates of the intersection of the isarithm and side. The side of exit, if not a border side or the starting side, becomes the input side into the next triangle, and the procedure continues.

An important aspect of this technique is that a simple logical search procedure is required at all stages; the method therefore lends itself to ease of programming. In addition, only the simplest form of interpolation need be performed, namely linear. A disadvantage of this technique is therefore that continuity of the orientation of the isarithms does not exist, as contiguous triangles do not generally form a continuous surface (eg. CIIS - Applied Research Cambridge, 1979b). If smooth isarithms are required, the string of interpolated points must be subjected to some continuous smoothing function, usually some form of piecewise cubic

polynomial function, generally a spline (see 2.4.1.2). However this may result in isarithms which cross in steep areas (e.g. Connelly, 1971).

2.4.3.1.2 Isarithm Generation Using Functions.

In order to overcome angular or crossing isarithms, a few methods have been evolved where interpolation occurs utilising a polynomial, the resultant isarithms being derived from the functions fitted to the triangulation. This results in smooth, distinct isarithms.

One such method has been implemented by McLain (1976). Essentially this is a hybrid method of interpolation, in which the final interpolation involves a weighted average of local patchwise functions. Assuming a triangulation net around each data point, local quadratic polynomials are fitted to the surrounding triangles, usually by weighted least squares. In the derivation of any point within a triangle, a weighted average of the three functions corresponding to each of the vertices is derived:

$z = w_1 f_1 + w_2 f_2 + w_3 f_3$;

where f_1, f_2, f_3 = the 6-term quadratic polynomial functions fitted at the three vertices to their nearest neighbours,

w_1, w_2, w_3 = the corresponding weights equalling zero on the triangle side opposite the relevant vertex (McLain, 1976).

To ensure smooth transition from one triangle to the next, it must only be ensured that each weight w and its leading derivatives are identically zero along the side of the triangle opposite to the vertex. McLain suggests that the weighting function be proportional to the distance cubed.

2.4.3.2 Isarithm Interpolation Utilising Grids.

Most isarithm interpolation programs require a regular grid either as input, or derived by interpolation. Indeed Connelly stated that,

> "there is now widespread agreement that the automated production of contour lines of a surface must proceed from a regular grid" (Connelly, 1971, 59).

The grid may be a more efficient form of point storage, but it creates several problems in the interpolation of isarithms. Most packages therefore derive a triangulation mesh from the grid as a first stage. This has the advantage, as previously shown, that it offers a unique solution.

Alternatively, isarithms may be generated directly from the grid, or a further interpolation function applied. As suggested earlier, linear isarithm interpolation in a grid leads to ambiguities which do not exist if a function, invariably a cubic polynomial, is utilised. The actual method of isarithm searching must also be considered. In both function fitting and direct isarithm threading, the isarithms are traced through the grid. Alternatively, the grid may be sequentially scanned, thus producing a raster series of line segments. The latter method although frequently used early in automation when storage was at a premium is - with the

exclusion of CONSYS (Cederquist, 1976) - now seldom used, as it produces angular isarithm segments.

2.4.3.2.1 Direct Isarithm Threading.

Direct isarithm threading in grids involves linear interpolation, utilising the same methodology as for triangulation-nets but with one exception. When a triangle is entered in triangular threading only one possible exit side occurs (Figure 2.14a). In contrast, when a grid cell is entered, three possible exit sides may occur, creating two possibly correct alternatives and a third impossible alternative (Figure 2.14b). This problem may seem unsurmountable, although in reality one of the diagonals if searched for an isarithm intersection point will indicate the true path.

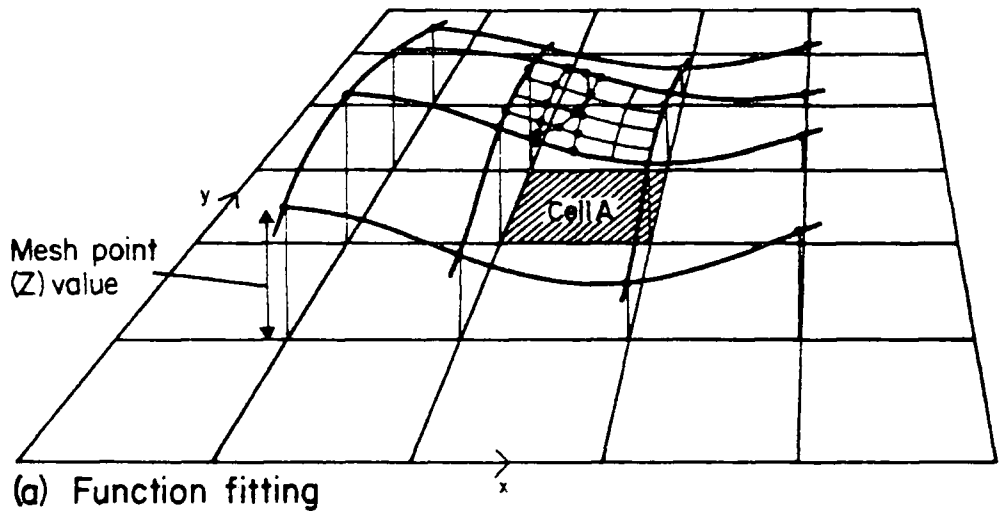
2.4.3.2.2 Isarithm Interpolation Utilising Functions.

Grid function isarithm interpolation is more common than direct grid and triangulation function isarithm interpolation. Commonly a bicubic polynomial is fitted by least squares to a local patch surrounding the grid cell under interest (Figure 2.15a). If a bicubic polynomial is used then, as 10 coefficients must be solved, a 4x4 node patch must be used. This results in an approximate fit with 6 redundant observations, thus generating extra smoothing and continuity between cells which may be encouraged further by using a spline fit. Having fitted the function, the isarithm may be systematically computed by solving the function in small

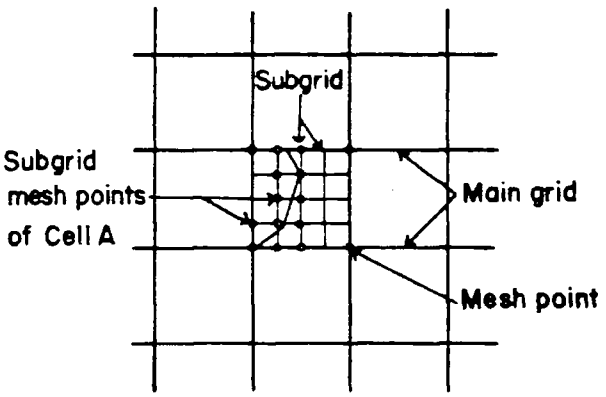
Figure 2.15

Grid function contouring

(After Calcomp,1971)



(a) Function fitting



(b) Sub-gridding

increments of x and y. A similar method is utilised in GPCP, where a finer sub-grid is generated within the grid cell (Figure 2.15b) and the denser sub-grid is solved by direct isarithm threading.

2.5 ISARITHM AND RANDOM-TO-GRID INTERPOLATION PROGRAMS USED

Several random-to-grid and isarithm interpolation programs have been utilised, although the two stages have where possible been kept distinct. Five programs were examined where both isarithm and random-to-grid interpolation were performed, namely GPCP; GINOSURF; SURFACE II GRAPHICS; SACM and SYMAP, although the SYMAP isarithm routine was not used as it produces line-printer output. Four programs were examined where only random-to-grid interpolation was performed, namely GHOT; BRAILE; MINCURV AND MULTI. Two programs were examined which generated isarithms from grids directly, namely GHOST and CONSYS. Finally DTC was examined which was the one direct random-to-isarithm program.

The programs and packages were not selected at random, but after careful consideration of several criteria (see Table 2.2). Clearly all available programs could not be used and therefore a subset of the populations had to be selected. The main consideration was that the subset had to be representative of the methods of random-to-grid and isarithm interpolation generally utilised. Following the literature search précised in the previous section, a short-list of desirable methods was established. This naturally corresponded

BASIS FOR SELECTION			PACKAGE
DESIRABLE	SPECIALISED	ADDITIONAL	
RANDOM-TO-GRID			
POINTWISE			
weighted average			SURFACE II GRAPHICS
		***	GPCP
search methods			SURFACE II GRAPHICS
weighting methods			SURFACE II GRAPHICS
		***	SYMAP
polynomial			GINOSURF
		***	GHOT
	Kriging		SURFACE II GRAPHICS
	Splining		MINCURV
PATCHWISE			
Polynomial			BRAILE
	Two-pass		SACM
HYBRID	Multiquadric		MULTI
GLOBAL			GPCP
ISARITHM			
DATA TRIANGULATION			DTC
GRID-TO-ISARITHM			
Triangulation			CONSYS
		***	GINOSURF
Rectangle			GHOST
		***	SURFACE II GRAPHICS
		***	SACM
Polynomial			GPCP

(*** - additional 'popular' packages selected)

Table 2.2 Criteria for the selection of interpolation programs considered in this research.

to the major categories discussed in 2.4. In addition, it was felt that some more specialised techniques might be worthy of investigation.

Having outlined suitable methods, it was necessary to find software to match these categories (see Table 2.2). Software were selected following consideration of certain criteria.

a. Frequently, appraisals of software tend to favour programs developed by the researcher. To avoid such a situation, a prime objective was to use (where possible) software that was extant, ideally an already established package.

b. Software utilised should be generally available (thus ruling out, for example, HIFI - Ebner, Hofmann-Wellenhof, Reiss and Steidler, 1980 and SCOP - Assmus, 1974). Additionally, preference would be given to that available in the academic environment as it would probably be generally accessible to users in many disciplines in addition to commercial users (thus ruling out, for example, MOSS). It was recognised however that, in reality, many commercial users employ packages unavailable to other commercial users and academic users due, for example, to cost (eg. CPS-1).

c. Following from 'a' is that standard 'off-the-shelf' packages were preferable (eg. GPCP and SURFACE II

GRAPHICS), although program listings available in recognised literature would be a tolerable alternative (e.g. MINCURV and BRAILE). A final alternative would be self-developed algorithms and software (e.g. DTC and MULTI).

d. A final consideration was that, if two popular standard packages were similar, yet employed marginally differing algorithms which would provide a viable examination, then they should be both considered (e.g. the GHOT and GINOSURF random-to-grid routines and the CONSYS and GINOSURF isarithm threading routines).

Obviously the exact selection procedure was therefore quite complex, with many points to be considered: the evaluation procedure necessitated the seeking out of available software and its use on a number of computers remote from the University of Durham, the implementation and de-bugging of the software on the local computer and the creation and testing of personally-written software. Overall, it is believed that the selection provides a reasonably representative sample of programs and techniques in widespread use in the academic environment.

It has been said that,

> "users of trend surface analysis need to guard against employing one of the many 'canned' computer programs that are now available without thoroughly understanding what the method accomplishes and how" (Whitten, 1975, 282).

The validity of this comment is not limited to trend

surface analysis programs, and therefore in order to understand the results and explain any problems in future chapters, it is necessary to consider the interpolation methods of each program in turn. Further details concerning the program implementation are given in Appendix 1.

2.5.1 GPCP - General Purpose Contouring Program

GPCP (Calcomp, 1973) is primarily an isarithmic package into which regular gridded data or random data may be input. If the latter are input, some form of random-to-grid interpolation is required, as isarithms may only be generated once a regular grid exists. There are two methods of random-to-grid interpolation in GPCP. The simpler method involves the generation of a global polynomial surface, the more complex method involves a two-stage pointwise technique.

Throughout the package, weighting is derived from the function

$w = 1/(1-r^2)$,

where r is the ratio of the distance of the data point from the interpolated point over the maximum distance to any point being considered for that interpolated point.

The pointwise interpolation involves the derivation of weighted means of values derived from linear surfaces fitted at each data point. Once the data have been input, and unless the local surface is specified by the user (up to tenth-order polynomial function), first-order polynomial surfaces are fitted at each data point by a weighted approximate fit, such that the surface is constrained to pass through the data

point. The only control over the fitting is obtained by the user-specified variable 'NBORSG' (default=8), which specifies the number of nearest data points - irrespective of direction - that will be involved in the fit. The second stage of the interpolation involves searching for the 'NBORST' (default=8) nearest neighbours to each grid node to be interpolated. For each node, having newly derived or specified the local data surfaces, the intersection of the relevant surfaces and the grid node is established, and a weighted mean taken. This involved procedure therefore allows for the derivation of heights outside the local and, indeed, the global range of the data.

The global interpolation function allows the user to generate a polynomial surface of up to and including tenth-order surfaces. The polynomial surface is generated by orthogonal polynomials, the order being specified by the user, and solved at the grid nodes. The resultant product will be much smoother than the primary interpolation method, and indeed it is suggested to be a data smoothing technique for noisy data (Calcomp, 1973).

Having established a grid, isarithms may be interpolated. The grid is first made denser by the patchwise fitting of a third-order polynomial surface to the 4x4 grid surrounding the cell under interest (Figure 2.15). Using the user-defined variable 'NREFS', (default=5), and the newly-generated third-order patchwise surface, the main grid cell, 2x2, is 'densified' to a NREFSxNREFS sub-grid, and the isarithms are threaded through this sub-grid using straight line segments

generated by linear interpolation between the grid nodes. It should be stressed that the sub-grid is only temporarily stored, and thus every time an isarithm is found to pass through a main grid cell, the sub-grid must be re-established, with a resultant increase in time.

2.5.2 GINOSURF

The input to GINOSURF may be a user-supplied mathematical function (which is solved for each node of a regular grid), a set of random data (which are then interpolated onto a regular grid), or a regular grid. Before isarithms may be generated, a regular grid is required. Thus generating isarithms from random data may be considered a two-stage process - the interpolation of a grid and then of isarithms. The routine RANCON combines the two stages in one call or, alternatively, the user may call RANGRD (to form a regular grid) and then GRDCON to perform the isarithm threading.

The random-to-grid interpolation is that described by Falconer (1971), and is essentially a pointwise, least squares, ratio-weighted, second-order fit.

To obtain the height of the surface at the grid node, only those data points within a circle of radius s are selected. The user selects a parameter LP (default=24), which is to be the average number of points within a search radius. The algorithm then derives the value s, such that on average there will be LP data points within the circle. If any circular patch now contains fewer than six points, then the radius of that patch only is increased until at least six points are

included. Six points are required in order to avoid singularity problems at the fitting stage.

An approximate least squares fit is now generated at each node. The data points are weighted by a factor w derived from the formula,

$$w = (s-d)^2/d$$

where d = the distance between the data point and the given point,

s = the average search distance as defined above.

A paraboloid is estimated by the fit. Since four parameters are required to specify a paraboloid, it is essential that the circle contains at least four data points if the least squares equations are to be non-singular. Provided s is kept constant, the surface defined by this interpolating technique will be continuous and have continuous first derivatives.

Having thus derived the grid or input a regular grid, isarithms may be interpolated by a call to RANCON, DRACON or GRDCON. The method of isarithm tracing is that of Heap and Pink (1969). Open isarithms are traced from one boundary to another and closed isarithms are traced back onto themselves.

Once the initial point is established by linear interpolation between the two grid nodes, each isarithm is traced through the relevant grid cells. When a grid cell is entered, the centre of the rectangle is interpolated by a direct average of the four surrounding nodes, and four triangles formed (Figure 2.13). The crossing point at the next edge of the triangle under interest is derived by linear

interpolation and thus the isarithm is traced through each triangle and through each rectangle. Once the isarithm has been completely traced, there is an option to smooth the isarithm using a piecewise cubic function (McConalogue, 1970).

2.5.3 SURFACE II GRAPHICS

SURFACE II GRAPHICS (Sampson, 1978) requires as input either gridded or random data. If random data are input, a choice of three random-to-grid interpolation methods (GRID, KRIGE, and TREND) may be used to generate a regular grid, which is required before isarithm interpolation may be undertaken. A feature of SURFACE II GRAPHICS is that if a grid node cannot be evaluated, it is set to a default value and isarithms are not threaded through the immediate area between the point and the neighbouring nodes.

2.5.3.1 GRID.

The 'GRID' routine is a pointwise distance-weighted averaging method of interpolation. It may be either one- or two-stage.

If the one-stage method is used, each grid value is estimated by a weighted average of the sample data points found in the specified search around that grid intersection. The grid value will thus be within the range of the data.

Alternatively, a method similar to that of the GPCP pointwise method may be used. The first stage involves fitting least squares, weighted, first-order surfaces at each data

point to the nearby data points found by the specified search routine. The surface is constrained to pass through the data point being evaluated. For each grid node being evaluated, a local search is made. A weighted average is then made of the solutions of the local data surfaces at the grid node. Overall four variables may be specified.

a. The interpolation may be one- or two-stage - the default is one-stage.

b. The weighting function to be used in both of the stages of interpolation, the options being:

$w = (1-(d/(1.1s))^2/(d(1.1s))^2$ (default),

$w = 1/d$,

$w = 1/d^2$,

$w = 1/d^4$,

$w = 1/d^6$,

where s = the distance from the grid node being interpolated to the most distant data point contributing to the estimation.

c. The search may be solely on the basis of nearest neighbours (Figure 2.10b). The number of nearest neighbours to be used (default=8) is specified with the minimum and maximum search radius that will be used. If the correct number of points is not found, the search fails and the grid node is set to the default value.

Alternatively, the search may be based on a quadrant search (Figure 2.10c). Within each quadrant, the search is performed as for the nearest neighbour, with the proviso that a specified number of quadrants (default=3) must have at least one data point.

Similarly, an octant search may be specified. Sampson (1978) suggests that octant searches are better for linear data, but tend to introduce more smoothing than quadrant searches which smooth more than a nearest neighbour search.

Finally, the search may be based on a circular area of specified size around the point of interest (Figure 2.10a). A 'VRADIUS' search finds all data points within a specified distance of the point of interest. A minimum (default=8) and maximum (default=12) number of points are specified, and the routine incrementally expands the search radius from the minimum to maximum user-specified radius of search until a number of points within the range specified is found.

d. If any of the searches fail, the point or node under consideration is set to a default value. Additionally, the searches may be used separately for either stage of the interpolation.

2.5.3.2 Kriging.

The Universal Kriging algorithm used in SURFACE II GRAPHICS was developed by Olea (1972 and 1974), and is a pointwise interpolation method.

> "In theory, no other method of grid generation can produce more accurate estimates of the form of a mapped surface. In practice, the effectiveness of kriging depends upon the proper selection of several parameters, including the slope of the semivariogram. However, even with naive estimates of these parameters, kriging will do no worse than arbitrarily estimating procedures such as those in 'GRID'" (Sampson, 1978, 111).

Kriging is explained elsewhere in this chapter and therefore only a brief outline of the implementation will be discussed here.

The algorithm requests user specification of the slope of the semivariogram within the standard neighbourhood and wide neighbourhood which defaults to 1. In addition, an OCTANT search is required with different parameters for the standard and wide neighbourhood searches.

2.5.3.3 Trend.

'TREND' may be used to fit a global least squares polynomial surface of up to sixth-order (default=3) to random data. The resultant coefficients are solved to produce a regular grid. Unfortunately, this option was not available in the implementation used.

Having thus generated or input a regular grid, straight isarithms may be generated directly from the grid by linear

interpolation between the grid nodes. These may be smoothed subsequently by piecewise Bessel interpolation within each grid cell. First, a line is calculated at the point of intersection of the isarithm line with the edge of a grid cell, so that the angles between the line and the isarithm segments are equal. If the slope of the isarithm exceeds the user-specified value, it is set to the specified value. Once the slopes are established at both end-points of a segment within a grid-cell, intermediate values can be calculated by a Bessel algorithm if smoothing is required.

2.5.4 SACM (Surface Approximations And Contour Mapping)

The SACM package, dedicated to surface interpolation and isarithm interpolation only, was conceived by Assiter (1967). SACM, in its standard form, encompasses four methods of generating a regular grid. The grid may be derived from user-supplied coefficients of a power series where up to tenth-order may be input. A global orthogonal polynomial, whose degree is in the range one to eight, may be generated by least squares. This may be similarly solved for each grid node. Alternatively, local functions in the form of linear combinations of the orthogonal polynomials may be generated. Finally - and the only interpolation method that was available in the implementation accessed - is the standard 'Numerical Approximation'.

The numerical approximation system, described by Batcha and Reese (1965) and Walters (1969), is a two-pass patchwise system. In the first pass, grid nodes of squares that contain

one or more data points are calculated using a patchwise least squares, weighted linear surface ('secant' plane) passing through the data point, or average point if there is more than one. The nearest data point in each of the octants surrounding the point is used in calculating this surface. The surface is then used to evaluate the four surrounding grid nodes. The second pass establishes the values for all remaining grid nodes by progressively working outward from the already calculated mesh points. The grid nodes are calculated by averaging two planes, the second of which is a tangent plane determined by means of a first-order finite-difference equation, which reflects the slopes of the planes at nearby calculated grid points.

Isarithm interpolation requires an interpolated regular grid, the isarithms being traced through the main grid by linearly interpolating across individual grid cells.

2.5.5 SYMAP (SYnographic MAPping Package)

SYMAP is perhaps the earliest and best known isarithm package, producing line-printer isarithm maps from a grid generated from random data. It should be stressed that the program does not allow for a regular grid to be input directly, hence input must consist of x,y, and z values, which - unlike the other commercial packages - must have their origin at the top left. The random-to-grid interpolation algorithm (Shepard, 1968) is pointwise and uses distance and direction weighting in a weighted mean of local data tangent planes.

Once the data have been input, local linear polynomial surfaces are fitted at each data point to the surrounding data points, and constrained to pass through the data point of interest. The number of surrounding points used varies with the search procedure, which is a similar technique to the SURFACE II GRAPHICS 'VRADIUS' search. An initial search radius and maximum search radius are specified by the user. Additionally, the minimum (default=4) and maximum (default=10) number of points to be searched for are also specified, with the proviso that an average seven points should be found.

For each grid node, the data points are scanned for points within the search radius specified, which is locally altered until the specified number of points is found. Using the resultant data points, the linear surfaces previously defined are solved at the grid node, and these values are used in the weighted mean. However, unlike many other packages, the intermediate values derived from the linear polynomial surfaces are restrained so that

> "the effect of slope will be limited to one-half the isarithm interval" (Shepard, 1968, 520).

The weighting factor encompasses both the node/data point distance and the directional dispersion of the data. The former is derived from a simpler function of distance squared. The latter, the directional dispersion, is derived as follows. Assuming two data points $D_1, (x_1, y_1)$ and $D_2, (x_2, y_2)$ and their distances to the grid node $P, (x, y)$ as d_1, d_2, the cosine of the angle may be defined,

$$\cos (D_1 P D_2) = ((x-x_1).(x-x_2)+(y-y_1).(y-y_2))/d_1 d_2 .$$

Since for all angles $-1 \leqslant \cos(\theta) \leqslant 1$,

where $\theta = \langle D_1 P D_2$,

it follows that if $t = 1 - \cos(\theta)$,

$$0 \leqslant t \leqslant 2.$$

Therefore, if the angle $D_1 P D_2$ is small, t will be near 0. As the angle increases to 180°, t increases to 2. t may therefore be used as a weighting estimate for angular dispersion. This weighting is combined with the distance weighting to obtain the general function,

$$w = d^2.(1 + t),$$

and the standard formula,

$$z = \left(\sum_{i=1}^{n} (w_i z_i)\right) / \left(\sum_{i=1}^{n} w_i\right) \text{ applied;}$$

where n = the number of points found by the searching routine.

Having generated the grid cells, the individual symbols on the final line printer map (if the relevant options are selected) are assigned in relation to the interpolated value and the results printed row by row. Thus isarithms as such are not generated in such a procedure. However, line-printer maps were not seriously considered in this research, although they do provide a useful 'quick-look' facility.

2.5.6 GHOT

GHOT is a random-to-grid program utilising the routines REGRID and DWLSQ1 (Worster, 1979). The interpolation method is that of a pointwise, distance-weighted, least squares, second-order polynomial fit. It is thus similar to the GINOSURF 'RANGRD' routine, with the exception that a full

6-term quadratic polynomial is used and the search and weighting functions are slightly different.

The search radius 's' is first computed from the full data set such that, on average, 30 points will lie inside a circle of this radius centred on each grid node. For each node in turn, all the data points are checked to see if they lie within the circle centred on the node, and if so, the weighting factor,

$$w = (s-d)^2/s^2,$$

where s = the radius as defined above,

is computed. At this stage if any point is found to lie within a minimal tolerance, the node value is set to the data value. If no data point lies within the required tolerance, the pointwise surface is fitted by least squares using all the data points within the search radius and solved at the grid node. Unlike most programs, no allowance is made for a lack of points within the search radius, and this creates serious problems of ill-conditioning if a poor distribution or low number of points exists.

2.5.7 BRAILE

BRAILE consists of a series of FORTRAN IV subroutines derived from various sources and modified by the author to perform random-to-grid interpolation. It computes a regular grid from randomly spaced data by the method of patchwise interpolation using local, low-order least squares polynomial fits. The grid is systematically sub-divided by user-specified variables into regular sub-grids. Each sub-grid is then used

as the centre for a rectangular patch, which is searched for any data points. This search patch may exactly abut onto neighbouring patches, or more likely overlaps may occur (see 2.4.1.2.2). The data points are then used in a least squares fit to produce local polynomial coefficients of user-specified order - usually first to fourth. At this stage, if there are not sufficient points to solve the order of polynomial requested or the number of redundant points is less than that requested by the user, a lower-order polynomial is fitted. The program terminates once all the sub-grids have been solved.

2.5.8 MINCURV

MINCURV performs random-to-grid interpolation using the difference equations derived from the requirement of minimum total curvature for the surface at the grid nodes. It is a pointwise method of interpolation, involving a two-stage process.

Throughout the program, a quadratic weighting function is used based on Falconer (1971),

$w = ((s-d)/d)^2$,

where s = the search radius.

The first stage involves the interpolation of a sparse initial grid using a weighted mean of the surrounding data points. The maximum search radius is defined by the user, and thus the process is similar to the SURFACE II GRAPHICS - 'GRID' interpolation (one-stage) with an area search (2.4.1.1.1).

Having established the initial grid, which is equivalent to every Lth element of the final grid (L=1,2,4,8,...), this grid

is further interpolated using the second stage process. The grid is 'densified' by halving the resolution, and the unknown grid nodes solved iteratively by systematically fitting pointwise, quadratic-weighted, fourth-order, finite difference surfaces to simulate splining. The iteration at any particular level of grid resolution only terminates once the difference between the grid nodes of the previous two iterations is less than DELTA - a user specified constant. The resolution is once again halved and the procedure continued until the required grid resolution and size are obtained. The method may thus be considered doubly iterative, first in terms of grid resolution and second in terms of precision of fit within any resolution.

2.5.9 MULTI

MULTI performs random-to-grid interpolation only, employing the theory of multiquadric analysis (see 2.4.1.2.2.1). It is thus a hybrid form of interpolation, involving a summation of surfaces. While several forms of quadric may be used, the cone (as described earlier) was the only form chosen.

2.5.10 GHOST

The GHOST package is a general graphics subroutine library of which the isarithm interpolation routines CONTIA, CONTIL, CONTRA and CONTRL represent a small section. While these routines only perform grid-to-isarithm interpolation, this may be from a regular or irregular grid. Irregular grids are

handled by routines CONTIA and CONTIL, regular grids by CONTRA and CONTRL. Similarly, calls of CONTRL and CONTIL will result in straight line isarithms, while CONTRA and CONTIA result in smooth continuous isarithms. Isarithms are generated by diagonally dividing each rectangular grid cell into two triangles, and using inverse linear interpolation to determine the intersections of the isarithms with the sides of the triangles (see 2.4.2.1 and 2.4.3.1). If linear isarithms are requested, the intersection points are connected by straight lines. Smooth isarithms utilise the curve fitting routine CURVEO.

CURVEO fits piecewise cubic splines to the intersection points thus ensuring continuity of slope, and then approximates these curves by straight line segments to a default resolution. This results in variable length isarithm segments depending on the curvature radius. Where the radius of curvature is large, segments will be long, and vice versa.

2.5.11 CONSYS

CONSYS (Cederquist, 1976) only allows for isarithm interpolation from a rectangular grid which has regular or irregular increments along each axis. The interpolation method produces raster isarithm segments with the scan in y and within that in x. In any grid cell, the centre is first derived as a mean of the four surrounding grid nodes (2.4.2.1.2). The corners and the centre are then considered to be connected with line segments which are in turn numbered 1 through 8. Each segment is then tested in turn to see if it

has an intersection with the relevant isarithm. Any resultant intersections are computed by linear interpolation along the line segment, and once the 8 segments in a cell have been examined, they are sorted to provide an isarithm segment within that grid cell. The rest of the cells are searched for all the isarithm levels and thus an isarithm map is built up cell by cell.

An inherent feature of CONSYS is thus the sharp linear isarithms, which require a plotter with high repeatability if the map is to look reasonable. This could be overcome by a second pass to link all line segments together; however CONSYS relies on simplicity of method and results.

2.5.12 DTC

The program DTC is the only random-to-isarithm program evaluated. It consists of a series of routines written by Green and Sibson (1978) which perform a Delaunay Triangulation, together with a series of routines which convert the raw output from the triangulation into a series of isarithms, written by the author.

A consideration of the triangulation methods available is given in 2.4.2.2. Sibson states that,

> "a triangulation is regarded as good for the purposes of interpolation if its triangles are nearly equiangular" Sibson (1978, 243).

Additionally, he proves that,

> "there is only one locally equiangular triangulation of the convex hull of a finite set of distinct data sites" (Sibson, 1978, 243),

and that that triangulation is the Delaunay triangulation,

the dual of the Dirichlet/Voronoi/Thiessen tesselation. Thus the Green and Sibson (1978) subroutines were adopted and adapted for isarithm interpolation.

Having read the data points, the program performs the triangulation which is then developed into an integrated triangulation network. Any isarithm levels may be requested by the user, and the resultant isarithms are threaded by linear interpolation. The external triangles are first searched for sides with a starting point at a particular level. Once an appropriate side is found, linear interpolation takes place to define the starting point of the isarithm. The other two sides of the triangle are searched for an exit point for the isarithm which becomes the entry point to the next triangle. Thus the isarithm is passed from triangle to triangle until it reaches another external side. The internal triangles are next searched and a similar procedure encountered to thread the isarithms.

Following the generation of the resultant straight-sectioned angular isarithms, a piecewise cubic polynomial may be fitted to the interpolated points on the isarithm to produce smooth isarithms. The smoothing routine is the algorithm by Akima (1970) which is called once the whole isarithm has been generated. The user can control the smoothness by selecting the number of intermediate points between the originally interpolated points.

2.6 SUMMARY

It is obvious from the previous discussion that digital

modelling has become the catalyst for many surface-related applications in many disciplines over the last decade. While this research is restricted to isarithm interpolation, the importance of random-to-grid interpolation and grid- or random-to-triangulation methods as the kernel of all digital modelling applications has been stressed. Therefore all users in those disciplines and applications requiring digital modelling necessarily require an understanding of the relative accuracies of isarithm and random-to-isarithm interpolation. Additionally, while the surfaces and data samples utilised vary quite considerably, their characteristics must be carefully assessed in conjunction with the interpolation. Finally, it has been shown that the interpolation programs examined represent a reasonably complete and representative cross-section of the interpolation methods generally available.

CHAPTER 3

DATA

3 DATA

3.1 INTRODUCTION

It is generally accepted that, within any study of surface reconstruction, the two main variables are the characteristics of the data and the method of surface reconstruction (or interpolation). The latter was considered in the previous chapter. The former, data characteristics, may be sub-divided into the inherent properties of the surface (considered in Chapter 4) and the data sampling characteristics.

Characteristics of the sampling process are important since sampling processes must be considered a form of data compression and, as such, are of primary importance if the fidelity of the data is to be maintained. Similarly, interpolation must be considered as a form of data densification which can only be successfully achieved if the data characteristics have been preserved.

When establishing a data sample, there are four associated considerations. The sample size (or density) is usually pre-defined: Leberl (1973) has suggested that a direct linear relationship exists between accuracy of reconstructed data and sampling density. Similarly, the sampling technique and the data surface are usually pre-defined, although they may directly control the accuracy of the data. Finally, the sampling pattern (or type of data set) is usually a matter of choice although it is generally related to the data

application, or more likely to the user's interpretation of the optimum sampling pattern for the required application and resource availability - sampling on a grid often gives rise to access problems in the field.

This chapter will briefly examine the methods of classification of data set type in an attempt to relate the data utilised within this research to 'reality'. In the process of this examination, it will propose a multi-disciplinary classification which is widely applicable. Additionally, it will investigate some of the data constraints applicable to this research, the method of data capture and the resultant data characteristics.

3.2 CLASSIFICATION OF DATA SET TYPE

The importance of data set type is stressed by several workers (e.g. Morrison, 1971 1974b; Ayeni, 1976 and Leberl, 1973). General discussion involves a variety of sampling classifications and therefore if research based on these data set types is to be compared, these classifications themselves must be compared and contrasted. Several classification systems, summarised in Table 3.1, will therefore be briefly examined and a multi-disciplinary classification proposed. In a subsequent section, the observed data will be related to this classification.

3.2.1 Summary Of Existing Data Set Type Classifications

a. The 'Cartographic Classification' proposed by Morrison (1970), although based on that previously demonstrated by

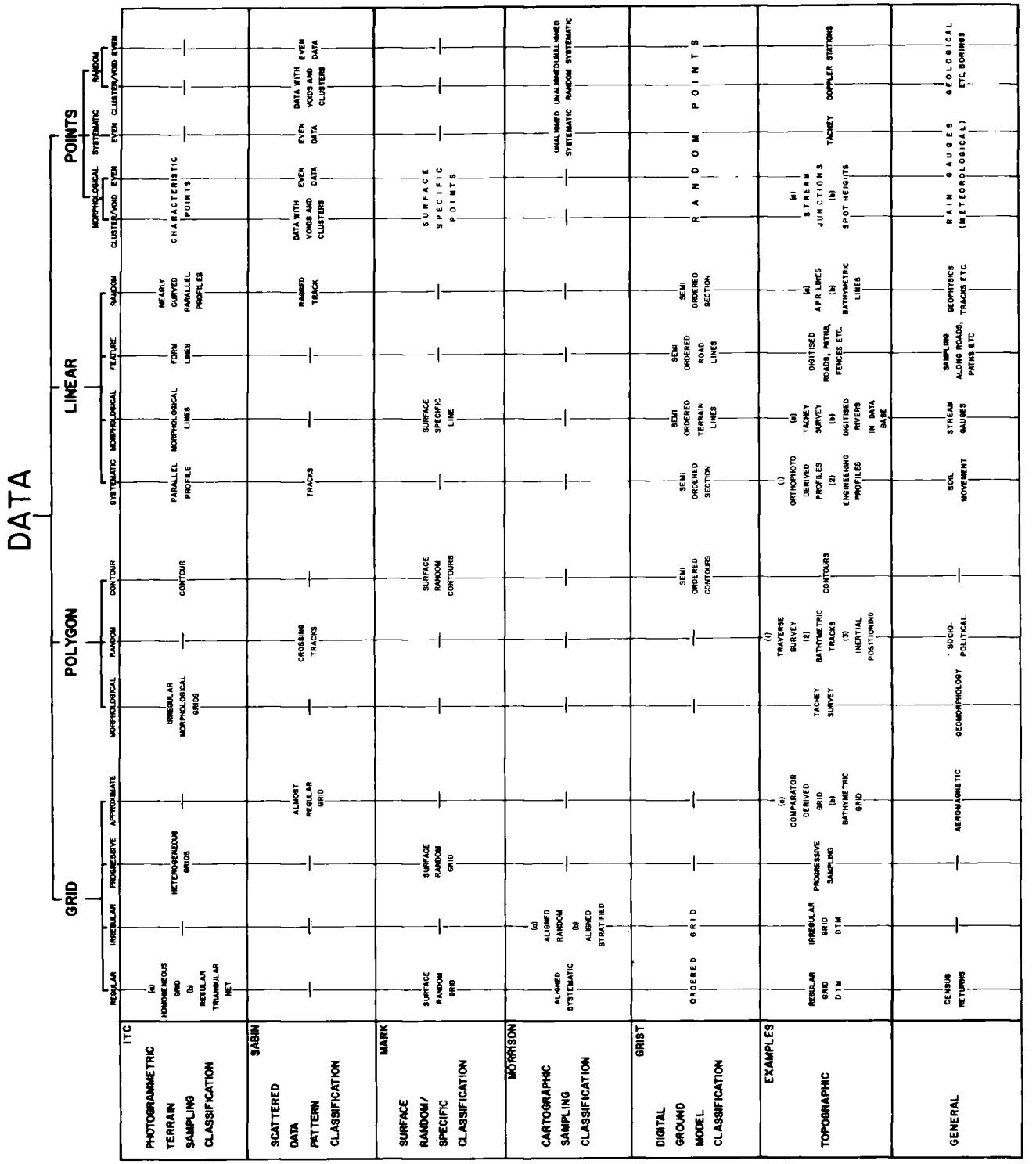


Table 3.1 Summary of data classification systems.

Quenouille (1949), is based on six sampling plans (augmented to seven by Ayeni, 1976) which he suggested could,

> "theoretically represent nearly all the possible sample point scatters in a plane" (Morrison, 1971, 26).

It implies no use of any prior knowledge of surface morphology and utilises a rigidly dictated random or systematic process of sampling. The aligned random and especially the unaligned random may result in clusters and areas of dispersion within the data. Alternatively, the unaligned systematic will encourage, and the aligned systematic will force, an even distribution.

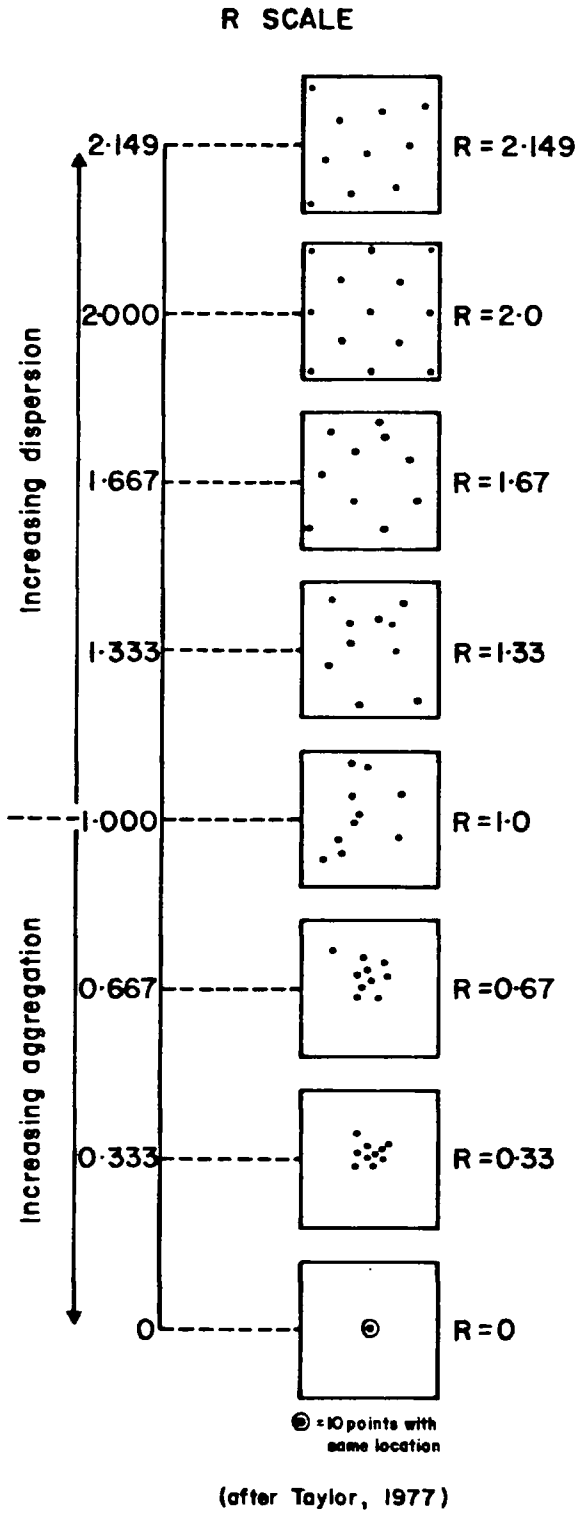
In conjunction with the pattern classification, Morrison promoted the concept of Nearest Neighbour Analysis (NNA) to quantify clustering. NNA provides a sliding scale of aggregation and dispersion of data (Figure 3.1) based on a ratio of average distance between points to expected average distance. Thus, the ratio expresses the divergence of the actual point pattern from randomness. A minimum score of 0 represents the situation where all points are superimposed, a score of 1 represents a truely random pattern, a score of 2, a square pattern and a maximum score of 2.149 results from a regular hexagonal (or equilaterally triangular) lattice.

A computer program (NN) was written for the derivation of NNA (see A1.4.15) and the Nearest Neighbour Statistic (NNS). This was used (see 3.4.5.2) to provide descriptors for the data sets used.

b. A classification of DTM samples is promoted by various workers in the ITC (e.g. Makarovic, 1973, 1976 and Braker,

Figure 3.1

NEAREST NEIGHBOUR ANALYSIS



(after Taylor, 1977)

1975). In general, it is either terrain-specific or systematic; no allowances are made for the purely stochastic sampling patterns which occur in most disciplines. This system is very much related to the way that a photogrammetric model is observed, and therefore the categories are not mutually exclusive for any data set.

c. Grist (1972) discussed a broader, and less precise classification than Morrison or ITC. Combining some terrain-specific reasoning with some systematic and random sampling, his ordered, semi-ordered and random point categories depend on the extent of pre-determination on the sampling. Grist fulfilled the true definitions of a DGM by relating potential interpolation methods to specific data set types.

d. Mark (1975a and 1979), in defining two distinct strategies for surface sampling, proposed a classification relatively similar to Grist, but without an interpolation/type relationship. The categories are quite loose and are related to the method of sampling (either surface-random or surface-specific) established a priori.

e. Sabin (1980) defined 'several patterns of scattered data' which are established independent of the data acquisition process. Additionally, like that of Morrison, it does not consider any inter-relationships in the third dimension.

3.2.2 Multi-disciplinary Classification

From the previous discussion and summary Table 3.1, it is apparent that most sampling classifications are either data acquisition-related or -independent sample descriptors. The Morrison and Sabin classifications only consider scattered data in two dimensions, and are based on the degree of statistical randomness of such data in the xy plane. In contrast, the ITC, Grist and Mark classifications emphasise the data properties inherent from the capture stage. None of the classifications are truly discipline- or situation-independent and therefore a more universal multi-disciplinary classification is proposed.

The classification proposed is the main classification in Table 3.1, where it is compared with the other methods already discussed. In the classification suggested, data are subdivided into grids, polygons, lines and points.

3.2.2.1 Grids

The use of grids in data sampling, with the exclusion of progressive grids (see below), produces perhaps the most objective, efficient method of data acquisition and storage. It allows fullest automation of data collection in most disciplines, although often encouraging redundancy. However, in some disciplines data acquisition may be highly inefficient, especially in getting to the sample points, for example geological bore holes.

Mark has argued that,

> "regular square grids, probably the most widely-used data structure for DTMs do not correspond with any specialists' groups approach, and thus do not represent an appropriate structure for digital terrain modelling" (Mark, 1979, 27).

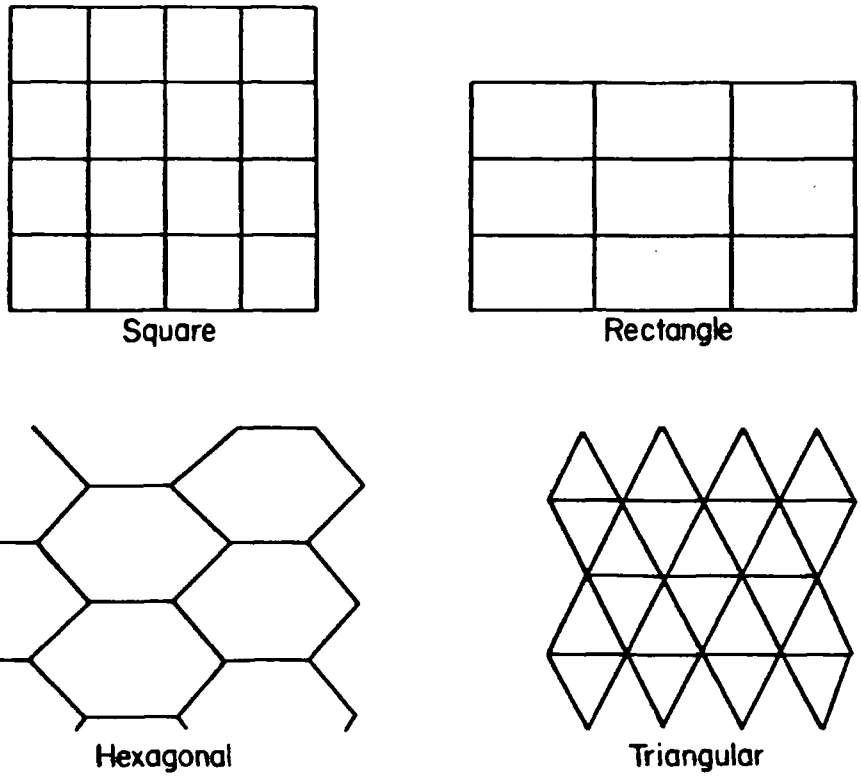
He suggests that geomorphologists would prefer continuous units, perhaps polygons bounded by breaks of slope; that surveyors would prefer triangulated irregular networks and that photogrammetrists would prefer contours and profiles. However, it has already been demonstrated (Chapter 2) that, since they represent the most convenient and perhaps efficient method of storage and processing given existing high level languages and array storage, they form the basis of most digital mapping techniques.

Assuming that a grid mesh is any regular polygonal shape (triangle, square, etc.), four commonly used grids may be defined (Figure 3.2a).

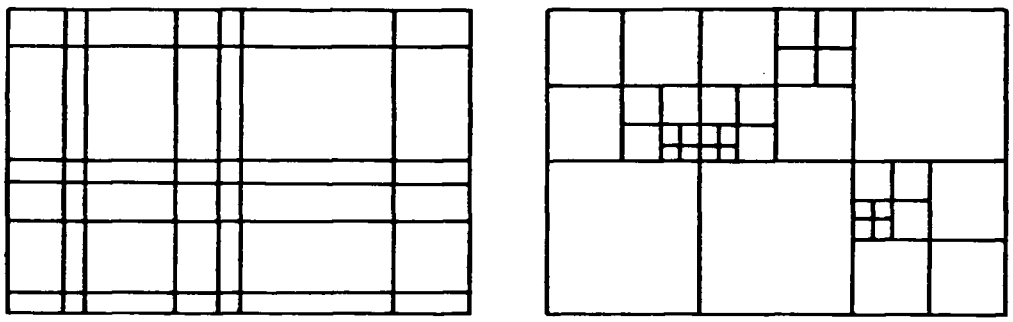
a. A regular grid is homogeneous and is pre-determined. Its success rests with its flexibility and efficient storage capacity, since the x and y co-ordinates of nodes are implicit in its definition. However the principal disadvantage is its tendency towards redundancy. The grid must be sufficiently dense throughout to portray the smallest objects within the area of interest. Tobler suggests that

> "if a function has no spectral components of frequency higher than W, then the value of the function is completely determined by a knowledge of its values at points spaced 0.5W apart" (Tobler, 1969, 243).

Hence, if the wavelength of smallest feature is 10m, then the grid resolution everywhere must be 5m, and considerable redundancy will occur in those parts of the

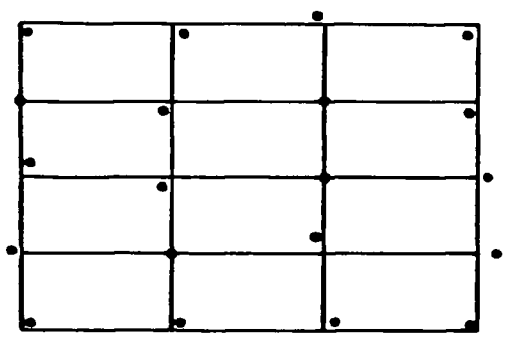


a) Regular grid (of differing mesh shapes)



b) Irregular grid

c) Progressive grid



d) Approximate grid

Figure 3.2 Commonly used grid types in digital modelling

data area where the wavelength of features is much larger.

The tendency has therefore been, where possible, to use either an irregular or progressive grid, both of which require some *a priori* knowledge of the surface being sampled.

b. *Irregular grid* sampling may remove much of the redundancy of the regular grid, but certainly not all. It may be generated by varying the resolution in each of the axes (Figure 3.2b) to define low wavelength features. However, per grid node, it is not quite as efficient as a regular grid for data storage as, in addition to the grid values, vectors containing the abscissae and ordinates must be stored.

c. *Progressive grid* sampling removes this extra grid redundancy totally as, instead of all values on the abscissa/ordinate being observed, only the areas of interest are sampled. However, it does involve a less simple data storage method as some form of hierarchical storage is required for the different levels of grid resolution. The basis of progressive grid sampling is to start with the lowest resolution of grid necessary to give a good general data coverage. This basic grid is then iteratively 'densified' by halving the grid resolution in

the areas where it is larger than half the terrain wavelength (Figure 3.2c).

d. An approximate grid occurs where the sampling procedure does not allow a precise, a priori location of the grid nodes, although the resultant grid is almost regular (Figure 3.2d).

3.2.2.2 Polygons

A polygon lattice is a very approximate form of grid which, for some appropriate purpose, has an irregularly shaped structure containing variably sized areas defined as separate entities during sampling. Included in this category are contours (in so far as inter-contour regions are polygons), morphological polygons and random polygons. Like all other data sampling techniques, they suffer in comparison with grids in that at least three values are required to be stored to define each point. However, this may be compacted to a limited extent by, for example, some form of node and segment storage system.

a. Contours, which can either be an a priori form of sampling or a post hoc data compaction or display mechanism, may be only directly sampled in a few surfaces (e.g. photogrammetry or by algebraic definition). This has generally limited their use although, without some method of data compaction, contours have a storage capacity

equivalent to a point sample. They may be sampled in terms of having Δx, Δy, Δx and Δy, or Δt constant (x and y are the cartesian axes and Δt is the time interval between digitising each point), although the method of sampling is largely dependent on the method of digitising, whether by automatic line following, manual line following, photogrammetric digitising or raster scanning.

b. Morphological polygons exist where the units are bounded by some physical phenomena (e.g. the crests and breaks of slope defining embankments or river valleys) and thus represent a surface-specific method of sampling. The polygon boundary must be of morphological significance within the surface, thus defining some homogeneous unit.

c. Surface-random polygons alternatively define heterogeneous units. They are a surface-random category which are formed from stochastic lines on the surface and may, for example, define property parcels or administrative areas.

3.2.2.3 Lines

Lines may be based on some a priori morphological knowledge of the surface, on some significant yet random surface feature, may follow a totally random process, adhere to some a

priori systematic process.

a. *Morphological lines* include breaks of slope, rivers, etc. with the proviso that the areas between morphological lines should be homogeneous. Breaks of slope are simpler to define than to generate as the practical definition is subjective.

b. *Feature lines* follow some other artificial but significant line on the surface. This may be a road, fence, etc. Strain (1972) used the term "special profiles" to define such features.

c. *Random topographic lines* form any other type of profile with random direction and variability. These may be inconsistent in orientation or direction, and are neither pre-determined nor morphologically significant.

d. *Systematic lines* are sampled *a priori* although they are surface-random. In general they are straight and probably parallel. It was this type of data sampling which formed the basis of the original DTM work by Miller and Laflame (1958).

In summary, lines may be sampled in most modes or combinations of modes (i.e. $\Delta x, \Delta y, \Delta z, \Delta xy, \Delta xz, \Delta yz, \Delta xyz$ or Δt -

where x,y,z are the cartesian directions, and t is time). Additionally, morphological lines may be sampled at some change of slope. Howard (1968) recommended that a resolution of 0.1mm at map scale along such lines is sufficient for cartographic reproduction, although 0.25-0.3mm is probably sufficient where curve fitting is used.

3.2.2.4 Points

Sampled points generally have little inter-relationship although there may be some overall factor controlling their selection. Selection may be based on morphological considerations, on some systematic sampling plan, or be totally random. Thus, the product will vary in distribution, with perhaps data clusters or voids, which may be described by the NNS.

a. Morphological points, following Mark (1975a), are selected in morphologically significant locations, i.e. peaks, pits, saddles and valleys.

b. Systematic points may be approximately selected by some pre-determined scheme, however the exact location is determined at the moment of sampling. This should result in an even distribution in x and y which will probably produce a z distribution comparable to the original surface distribution.

c. Random points are located purely arbitrarily at the time of sampling and should have no pre-determined correlation with any characteristic of the surface.

3.3 DETERMINATION OF DATA SAMPLING CRITERIA

It has been argued earlier that there are four criteria to be defined within any data sample. Definition of the data surface and the sampling technique relate to the source of the data and which method of observation was used to procure the sample. The size of the sample and the method of sampling must also be defined, either before or as the sample is collected.

3.3.1 Data Surface And Sampling Technique

Previous isarithmic and interpolation evaluations have been based on two approaches. In the majority of studies, no 'correct' or reference set of values has been observed and therefore either the results only show the range in performances that are attainable (e.g. Rhind, 1971 and Walters, 1969) or a reference is computed using one of the packages, and measures of deviation from an interpolation reference are produced (e.g. Walden, 1972 and Braile, 1978). Alternatively, a reference may be observed usually at the data collection stage, although this requires that a totally measurable surface be used. This has proved restrictive since only a few surfaces are feasible - primarily mathematical and

photogrammetric surfaces. The former have been used, for example, by Walden (1972) and Morrison (1970, 1971, 1974b), while the latter have had limited use, restricted to topographic surfaces, for example, by Ackerman (1978) and Assmus (1974, 1976).

In considering the limitations of these previous studies, various conclusions were drawn with regard to the nature of the data surface and sampling technique required:

a. Use of a computed reference surface was considered undesirable. While it may allow the establishment of measures of relative error between methods, it was hoped to be able to derive estimates of absolute topographic error for specific interpolation methods and point samples. In addition, it was hoped to relate interpolation error to point distribution. Clearly a reference surface derived from one type of point distribution would bias subsequent results and remove the possibility of establishing an unbiassed measure of absolute error. It was therefore considered necessary to utilise a surface which was totally measurable and a sampling technique which would allow measurement of any point on that surface. This would ensure that all data and any necessary references (grids and isarithms) could be easily observed.

b. Measurements would have to be of high accuracy. While some disciplines and interpolation methods frequently utilise relatively low accuracy data (eg. geophysical

mapping), high accuracy data would be sought since they would produce high accuracy statistics and could be of relevance to all disciplines.

c. The sampling technique should be capable of measuring a range of surfaces. There is no point in considering only one surface, as error might be related to quirks of the data. Several surfaces would therefore be observed.

d. Associated with c., a range of surface types were required since it was hoped to relate error of the interpolation process to surface characteristics. In order to make the results relevant to many disciplines, it was unsatisfactory to have all surfaces as smooth as those in Figure 2.3; as has already been suggested (2.2) a wide range of surface roughnesses are present in data used in digital modelling.

e. Finally, it has been demonstrated that a wide variety of data samples and sampling methods are already used in digital modelling (see 2.2 and 3.2). In order that the research would reflect this variety, the sampling technique/surface would have to allow surface specific/random sampling and the procurement of grids, lines, polygons (contours) and points.

It is possible that some surface types and measurement

systems could have provided the data used within this research. However, as a result of the literature search and of experience, it was concluded that only mathematically-defined or topographic surfaces certainly met all the desiderata. It was further decided to utilise only topographic surfaces and to sample the surfaces using photogrammetry. This decision was based on the following:

a. This study has concentrated primarily upon the applications within DTMs. It is therefore appropriate to use associated surfaces, i.e. topographic surfaces.

b. The author's experience lies in mapping the topographic surface and therefore problems and limitations of the measurement and interpolation process would be better understood.

c. Everyone is familar with the topographic landscape and therefore should be able to appreciate readily any interpolation results.

d. Many mathematical surfaces, while produced from apparently complex formulae, are inherently highly autocorrelated. Morrison has been the main advocate of mathematical surfaces, and yet his most complex surfaces are certainly highly autocorrelated (Figure 2.3). Indeed Walden, who also used mathematical surfaces in her study,

commenting on Morrison's choice, states,

> "Selection of these smooth test surfaces restricts Morrison's conclusions to the ability of his interpolation models to reproduce mathematically defined functions, which are not representative of topographic or other naturally occurring surfaces" (Walden, 1972, 52).

In contrast, most users of isarithmic and interpolation packages are interested in modelling natural surfaces, which often contain sharp discontinuities as well as areas of high and low autocorrelation.

e. Use of mathematical surfaces involves problems of determining surface-related samples. Only statistically random points, generated by some random number generator or grids, have been utilised in the known studies. In contrast, when using a photogrammetrically-derived surface, a visual model is being measured. The observer can relate to this model more easily and thus observe surface-specific points in addition to random or grid points. Similar points could be generated from a mathematical surface but would require multiple pains-taking solutions of equations, or by visual inspection of and interpretation from contoured maps or block diagrams.

f. The photogrammetric system simplifies the selection of purely objective sampling - grids and isarithms - since most plotters generally have clamping and slow-motion in all three axes.

g. Photogrammetry can give very precise measurements in a reasonably short time thus having benefits over, for example, ground survey.

h. The photograph used in the photogrammetric plotter records an instant in time. A unique solution is used in orientation before observation, which allows the model to be re-established and the unique surface re-measured if required.

Photogrammetrically-derived topographic surfaces thus represent a satisfactory solution to all the problems discussed earlier, and therefore must be considered at least as good as any alternative. Perhaps the only possible disadvantage is the possibility of error due to poor model control although, since relative and not absolute co-ordinates were required, then (providing the relative orientation was good) suitable accuracy would be obtained. Photogrammetry was thus chosen as the method of data acquisition.

Three topographic surfaces (models) of contrasting complexity were selected for the research and several data sets were derived for each. Descriptive statistics for each surface/model are summarised in Table 3.2 and block diagrams of each surface are shown in Figure 3.3. Each data set is displayed in Figure 3.4 with the relevant photogrammetrically-derived contours for the particular model.

The 'INCH' data sets were derived from a 1500m square area beside Inchnadamph near Loch Assynt. The area (Figure 3.3a),

DATASETS

	FORV				LIAN					INCH				
MODEL SCALE	1 : 5000				1 : 15,000					1 : 15,000				
MAP SCALE	1 : 5000				1 : 7,500					1 : 7,500				
RANGE IN X	60 - 700				0 - 750					0 - 1500				
RANGE IN Y	60 - 700				0 - 900					0 - 1500				
RANGE IN Z	27 - 57				50 - 150					50 - 250				
CONTOUR INTERVAL	2M				5M					10M				
GRID SIZE	65 x 65				51 x 61					51 x 51				
RESOLUTION	10M (2mm at model)				15M (1mm at model)					30M (2mm at model)				
	BREAKLINE	CONTOURS	GRID	RANDOM SCATTER	BREAKLINE	CONTOURS	GRID	RANDOM SCATTER	RIVERS	BREAKLINE	CONTOURS	GRID	RANDOM SCATTER	RIVERS
OBSERVED														
NUMBER OF POINTS	2229	9546	4225	250	579	5443	3111	250	291	1089	11107	2601	250	280
TIME TO DIGITISE(HRS)	7·25	8·75+6·50	24·00	1·00	2·25	3·75+3·00	12·50	1·00	1·50	3·75	6·25+4·75	12·25	1·00	2·25
DERIVED DATASETS														
NAME	FB	FC	FG	FS	LB	LC	LG	LS	LR	IB	IC	IG	IS	IR
TOTAL NO. OF POINTS	306	306	306	306	305	316	304	294	322	338	290	324	290	294
NO. OF INTERNAL PNTS.	250	250	250	250	255	271	238	250	265	266	247	257	249	243
N.N.A.	1·224	1·458	2·287	1·279	1·050	1·340	1·868	1·099	0·768	1·098	1·251	2·000	1·163	0·804

Table 3.2 Summary of data sets.

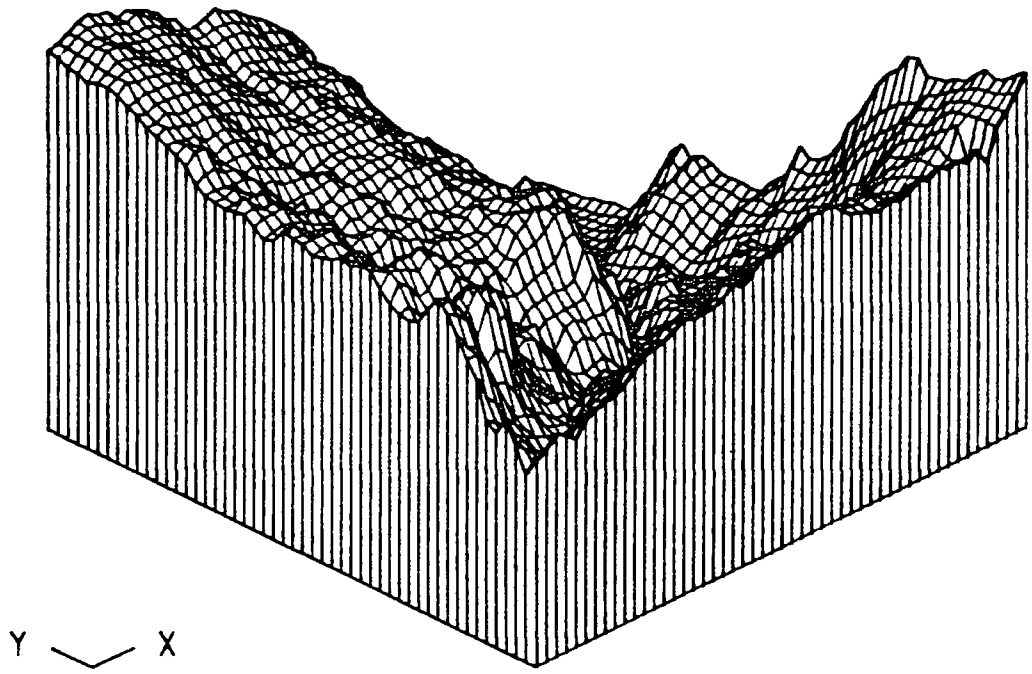
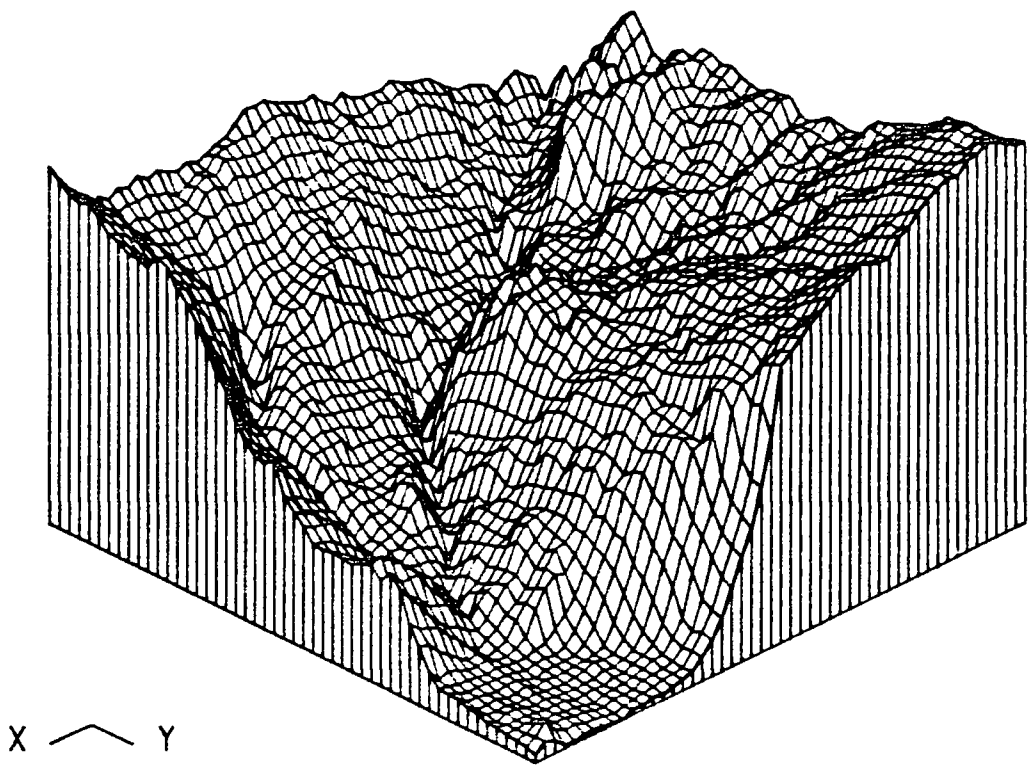


Figure 3.3a Block diagrams of INCH



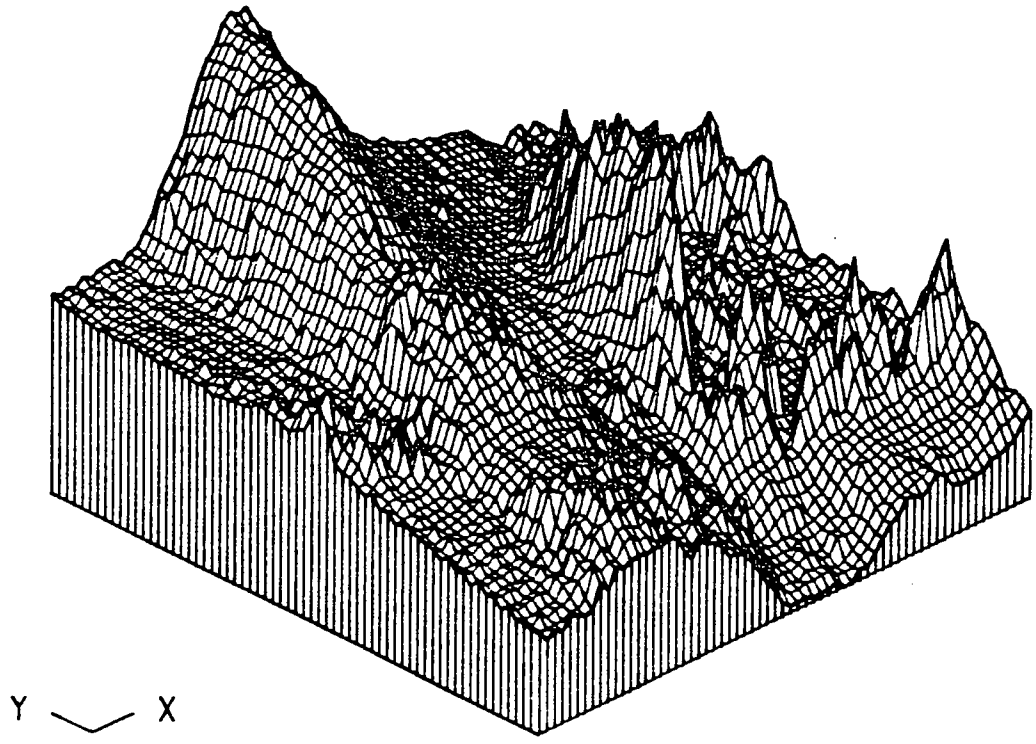


Figure 3.3b Block diagram of FORV

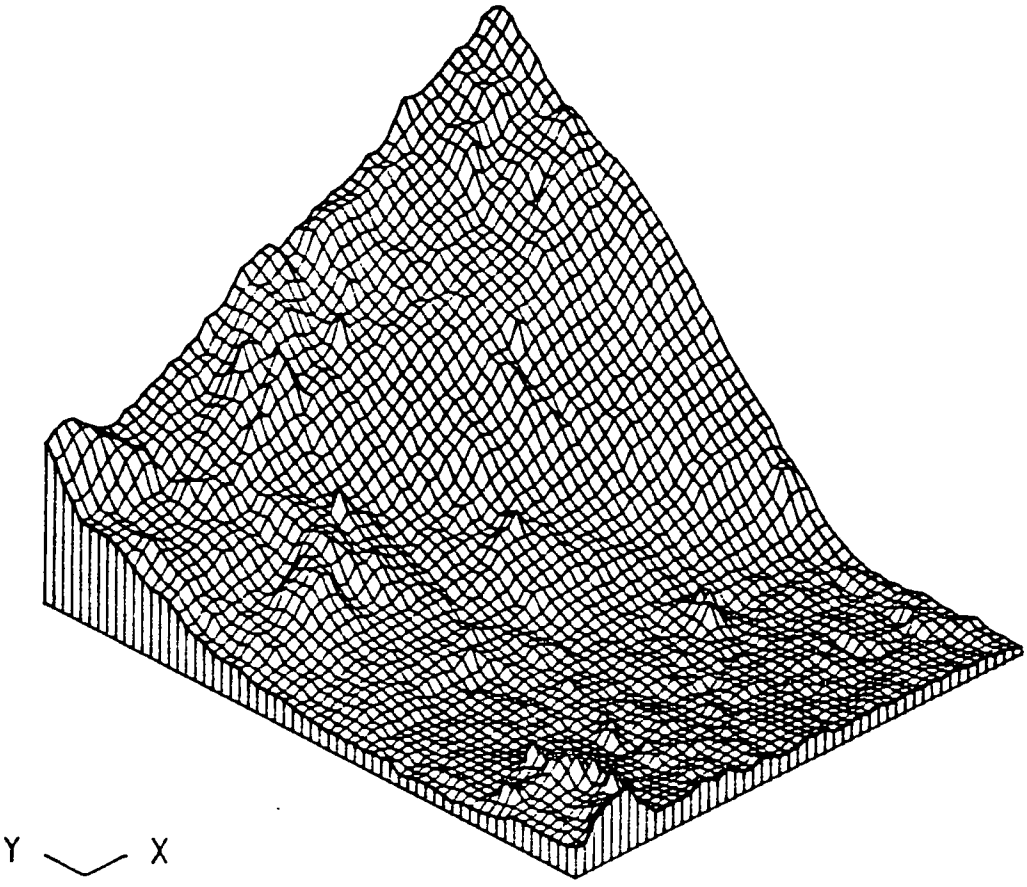
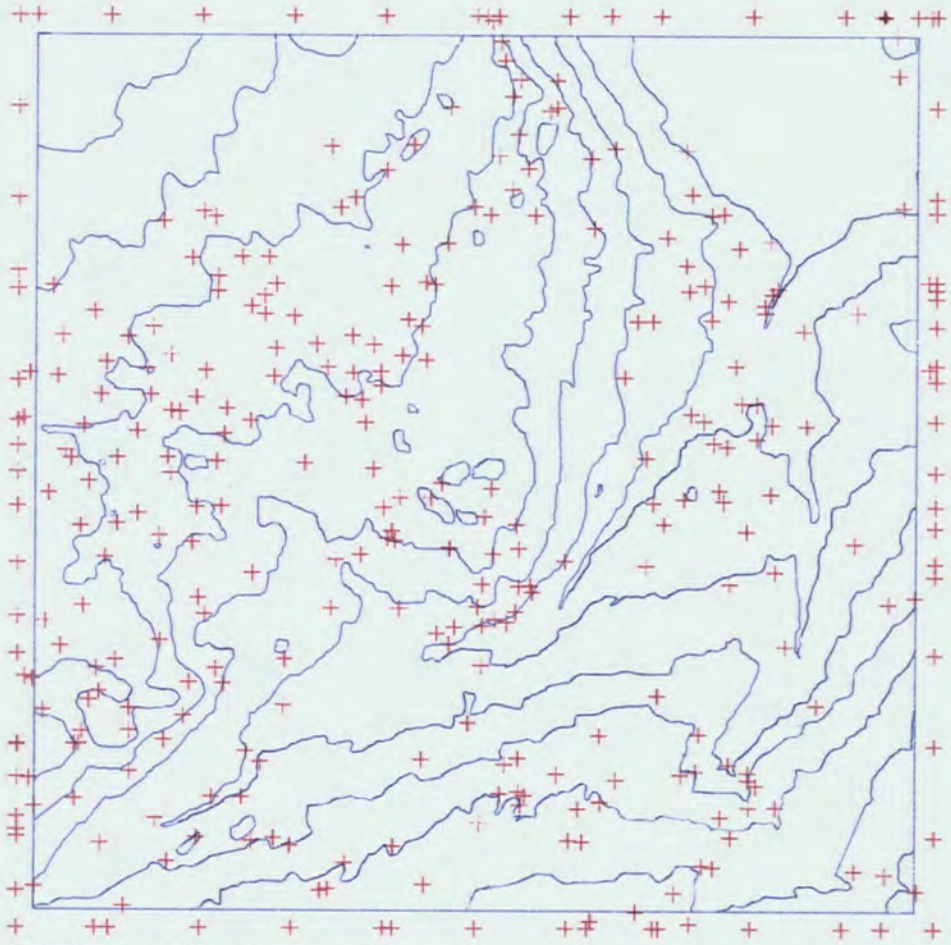


Figure 3.3b Block diagram of LIAN

Figure 3.4a



INCH-BREAKLINES

INCH-CONTOURS

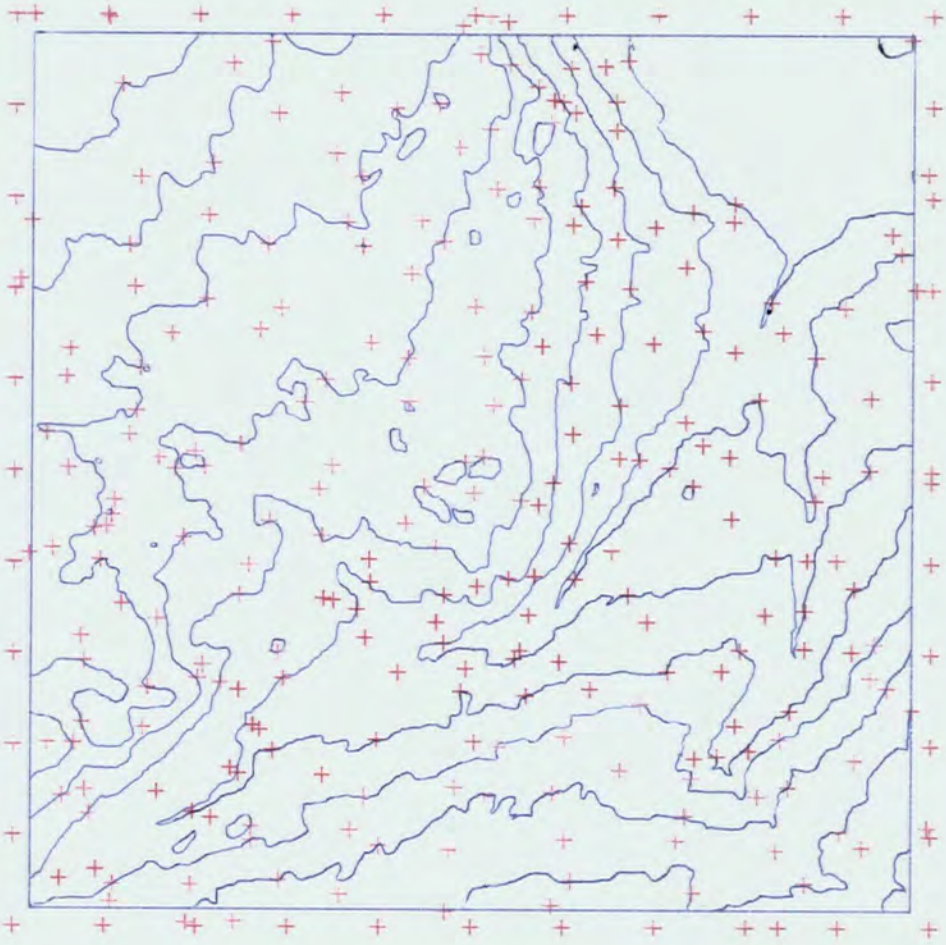
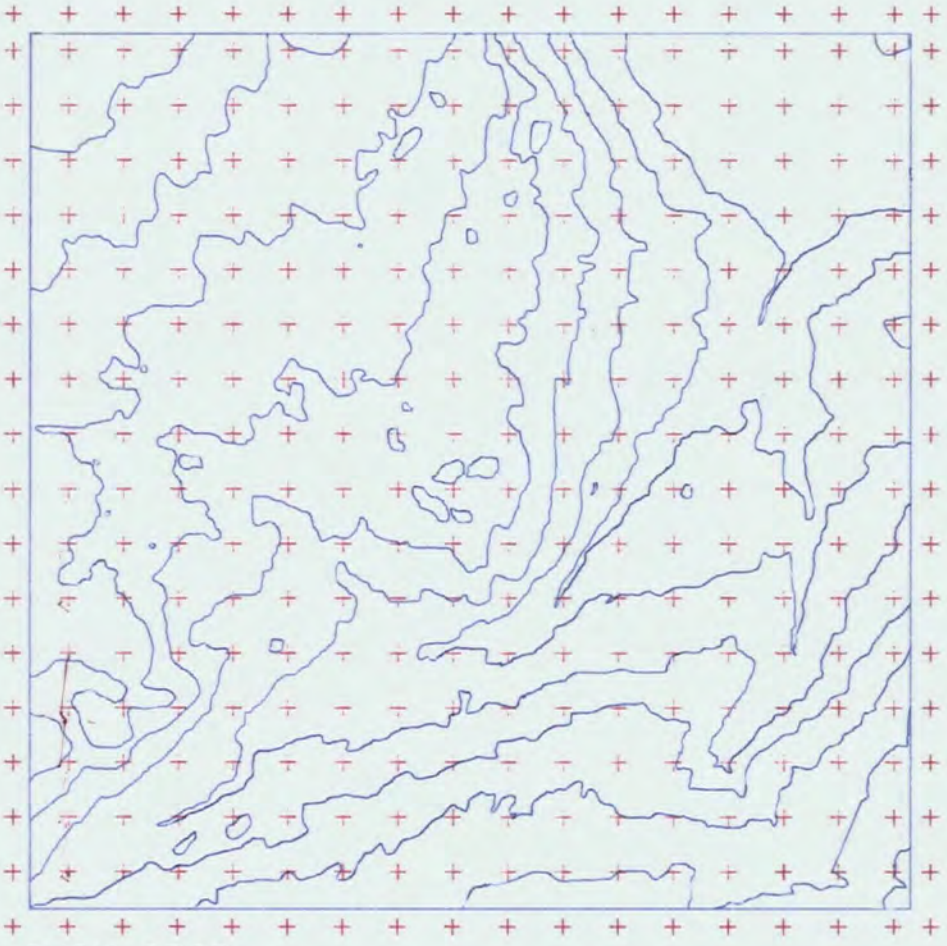


Figure 3.4b



INCH-GRID

INCH-SCATTER

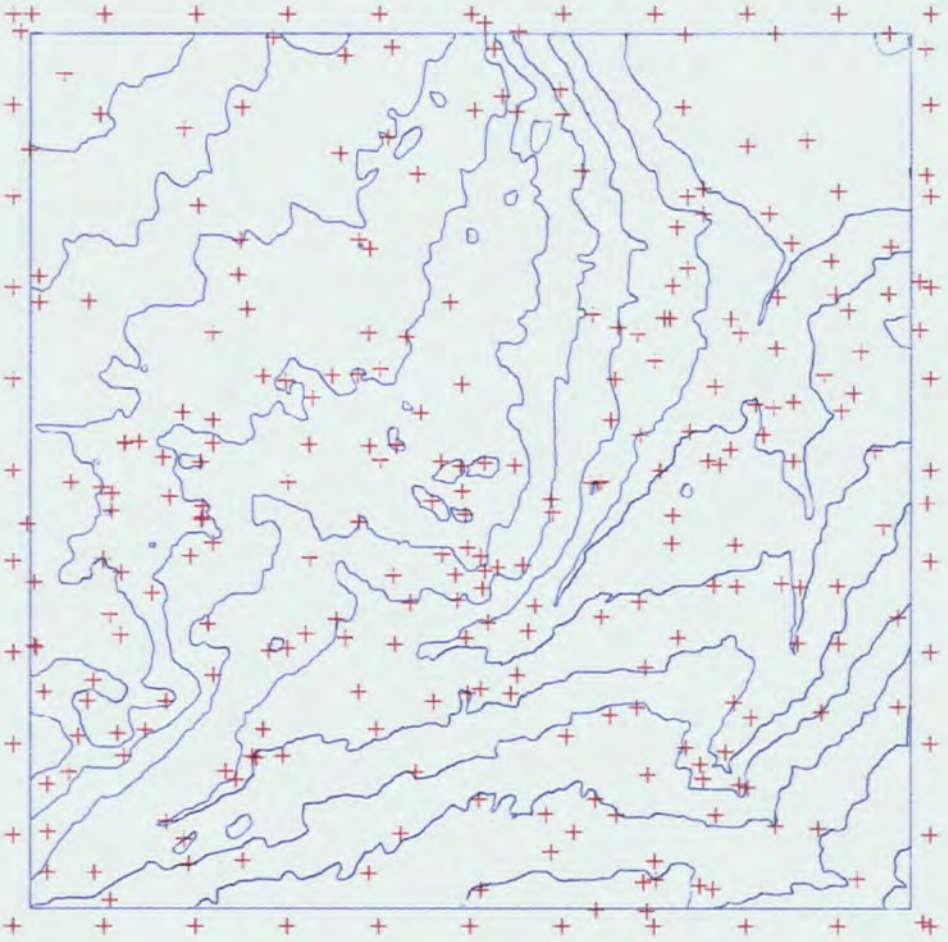
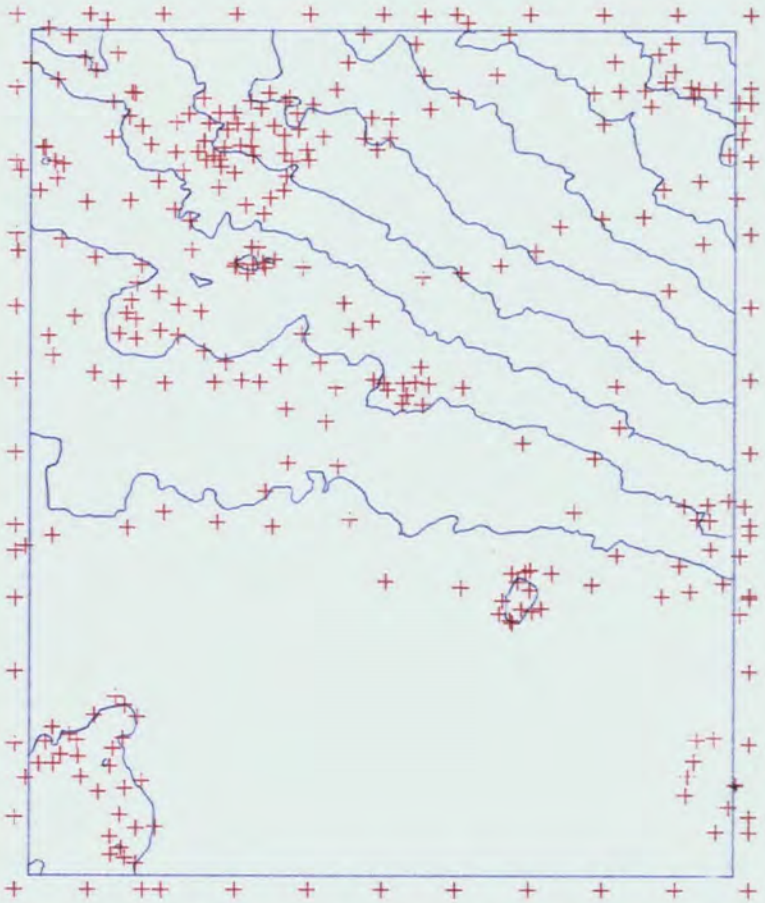


Figure 3.4c



LIAN-BREAKLINES

LIAN-CONTOURS

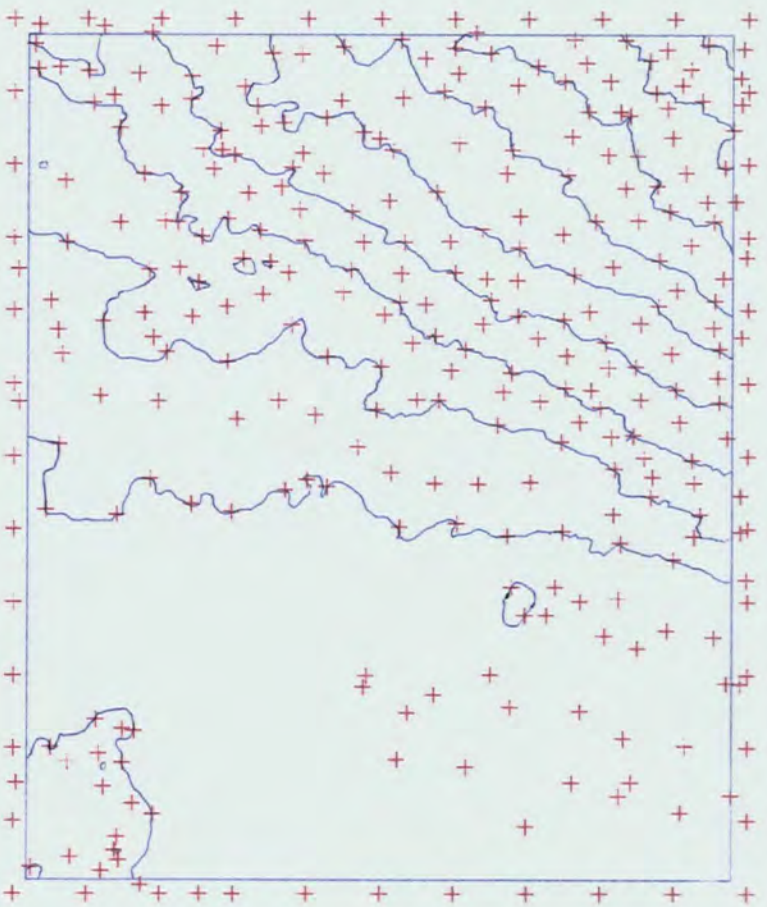
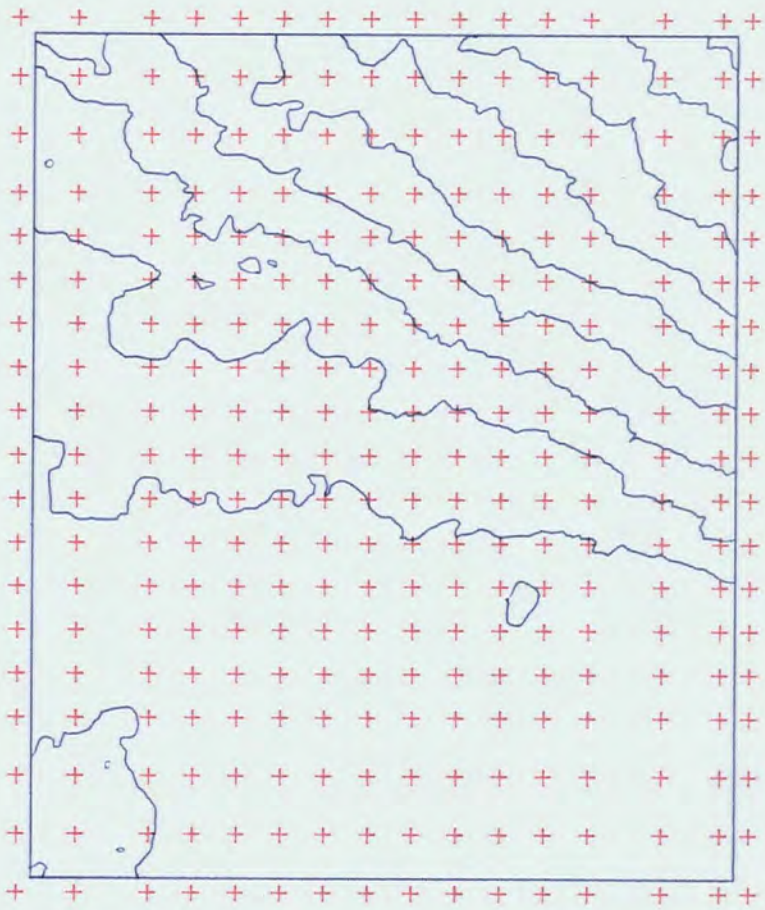


Figure 3.4d



LIAN-GRID

LIAN-SCATTER

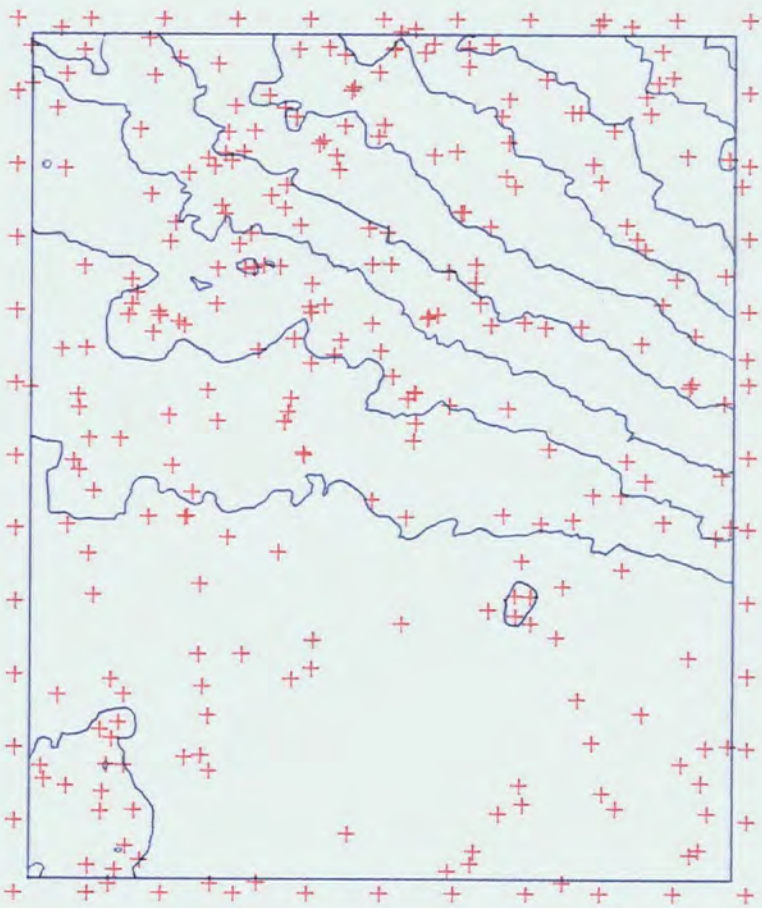
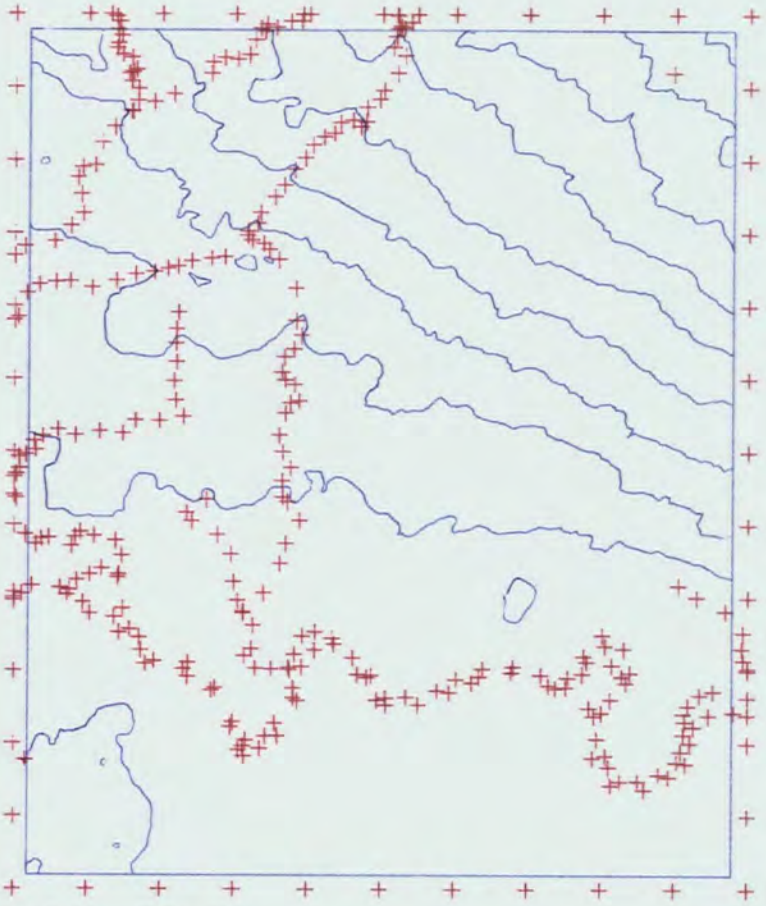


Figure 3.4e



LIAN-RIVER

INCH-RIVER

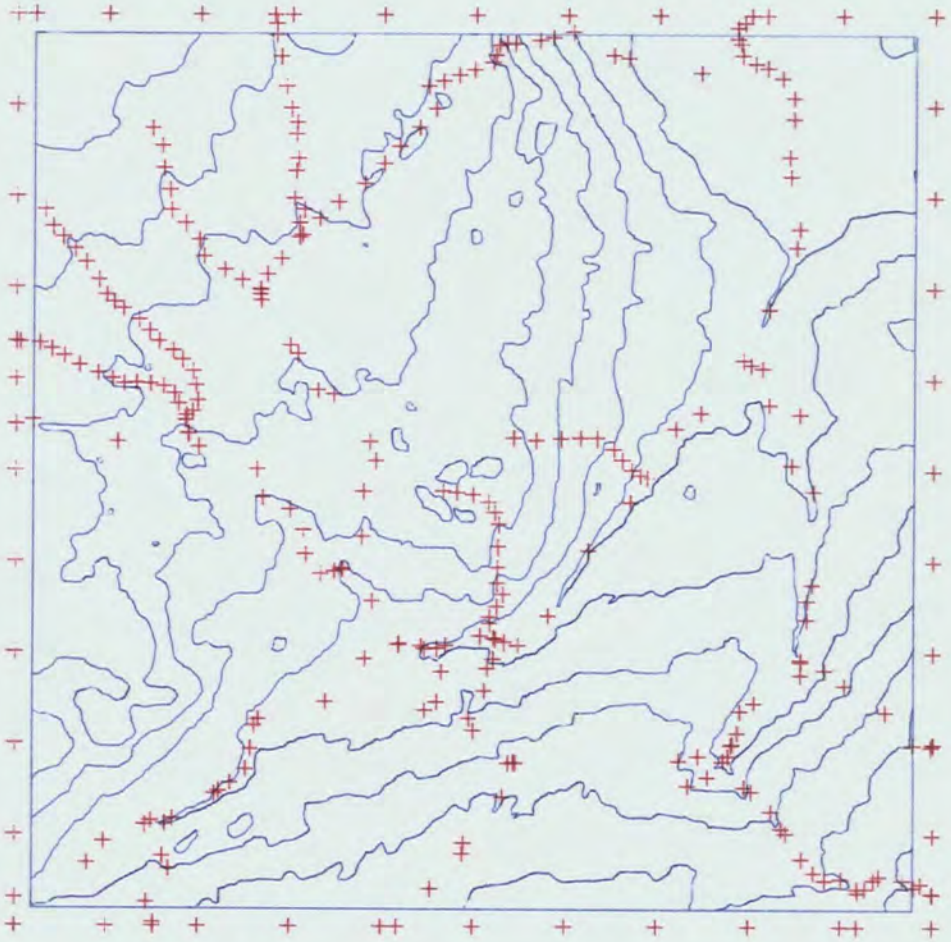
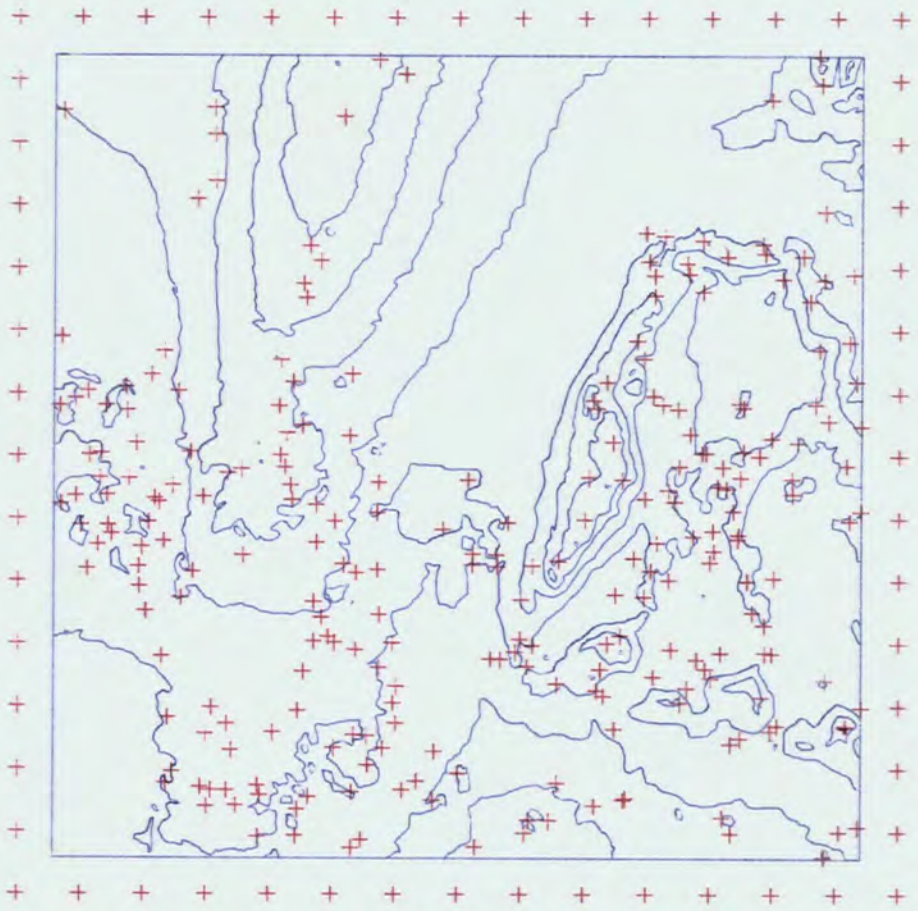


Figure 3.4f



FORV-BREAKLINES

FORV-CONTOURS

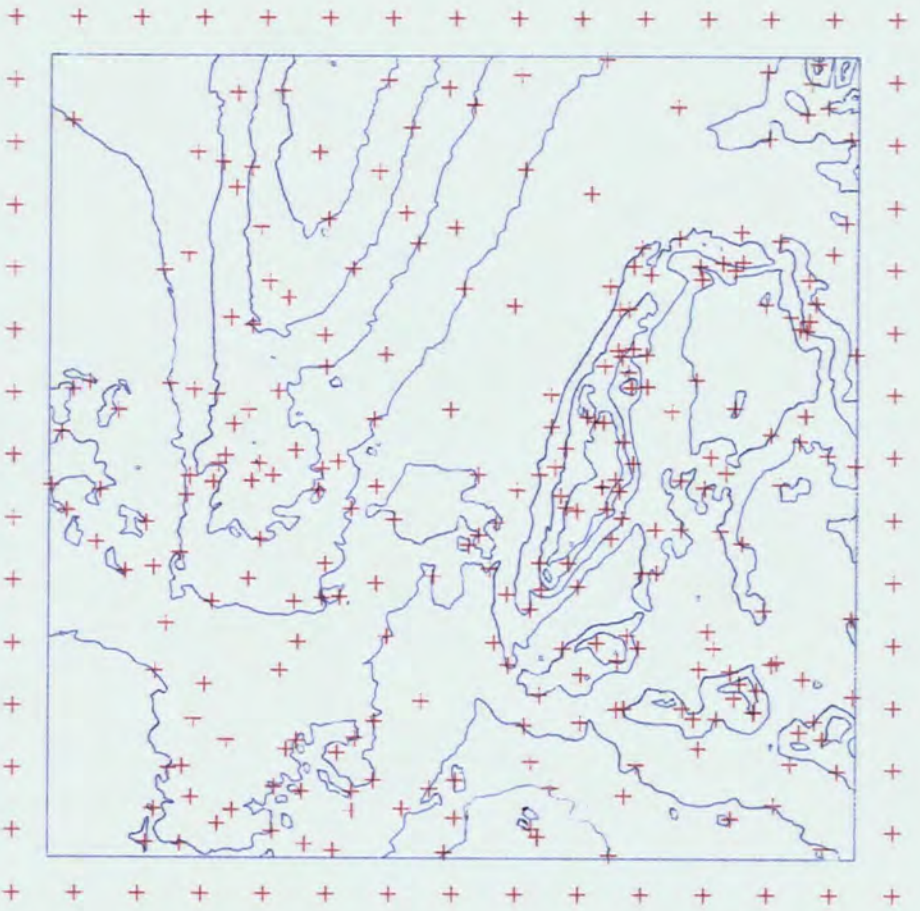
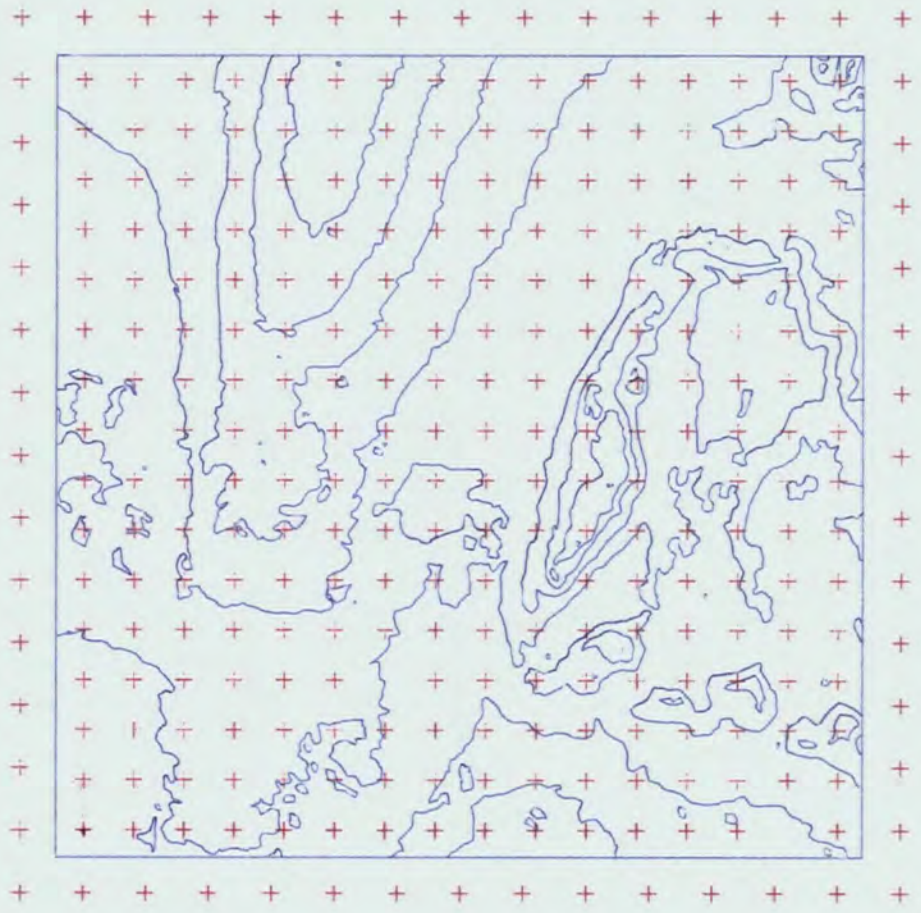
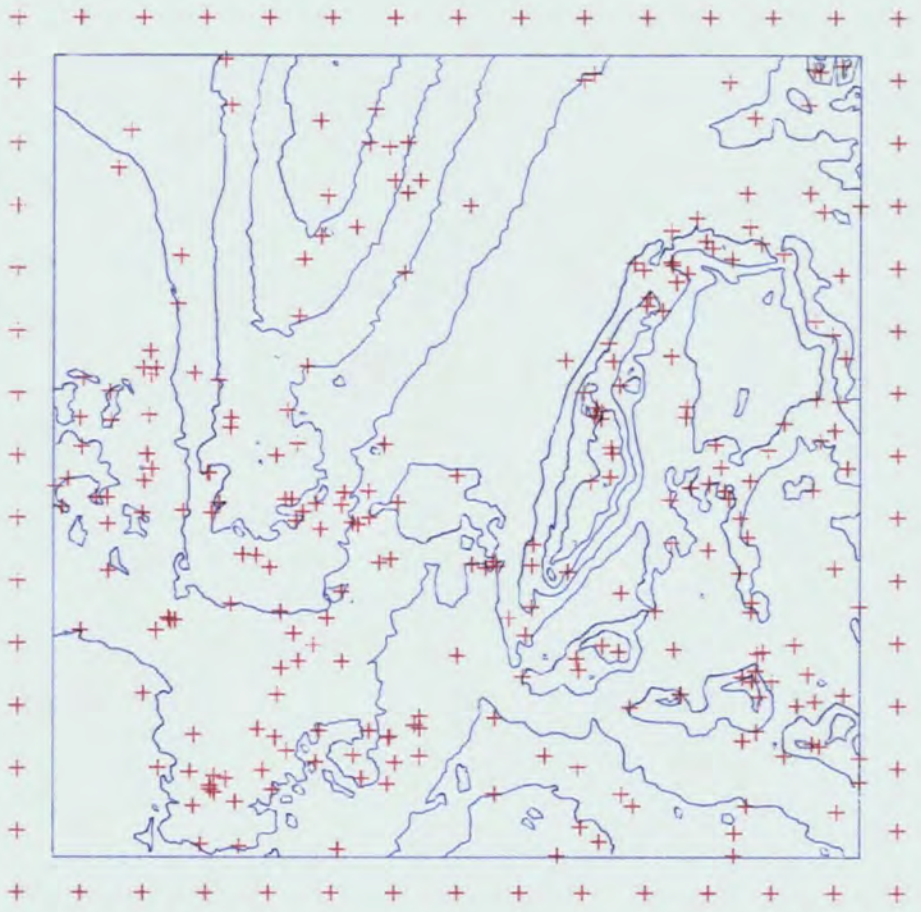


Figure 3.4g



FORV-GRID

FORV-SCATTER



which represents rough macro-relief, contains a dissected escarpment - consisting of a section of steep scarp slope, two sharply incised valleys and an undulating dip-slope.

The 'FORV' data sets were selected from an area of sand-dunes in the Sands of Forvie on the Ythan estuary near Aberdeen. The model area (Figure 3.3b) is 640m square and contains two major features, a large transverse dune and a large blow-out, in addition to several small dunes and blow-outs. The area was considered to represent rough micro-relief.

The 'LIAN' data sets were observed from a 750x900m area below the previous escarpment near Inchnadamph. The area (Figure 3.3c) includes a section of the River Loanan valley and the gentle, smoothly sloping Leathad Lianach, and may reasonably be considered a smooth surface.

The characteristics of all three model areas will be studied in more detail in Chapter 4.

3.3.2 Sampling Pattern

Five different sampling patterns were selected to be representative generally of the sampling classification described earlier, and therefore be representative of multi-disciplinary data. In addition, they were selected because they represented features that might be easily digitised from a standard topographic map sheet. The five sampling patterns chosen were as follows.

a. Contours. These were required as a reference to

enable a thorough examination of the contouring packages. However, it was decided to derive a subset of the contour points to test surface reconstruction from this polygon form of data.

b. Regular Grid. This was similarly required to examine the accuracy of random-to-grid interpolation. In order to study the effect of grid resolution on grid-to-isarithm interpolation, it was decided to observe a fairly large grid which could be subsequently halved and quartered and still produce a reasonably sized low resolution grid. A grid size of 12x12 was considered essential for the lowest resolution and therefore the full grid size would have to be at least 48x48. The full grid would also be re-sampled to provide a grid data set to study the effect of random-to-grid interpolation on this regular grid form of data.

c. Breaklines. These were observed as a form of linear morphological data to test random-to-grid interpolation. Breaklines are a common feature of most natural surfaces, and therefore would be used in most disciplines. Breaklines were selected such that they did not include any river points.

d. Rivers. These lineated features were also observed

since, whilst representing a similar type of feature to breaklines, they occur on standard topographic maps and are actual features on the ground which can be observed more objectively. Since no rivers existed in the FORV model, this data set is only applicable to LIAN and INCH.

e. Scatter. A scatter of points, representing truly random point data, was also observed, since this is typical of many data sets that are input to contouring routines. While this would not be generated by any statistically random process, it was hoped that by random movement of the plotting machine controls, a suitable sample could be generated.

3.3.3 Sample Size

In previous studies of random-to-grid interpolation, sample sizes ranging from 25 points to 400 points have been selected. Walden states that,

> "In general increases in sample size (100, 250, 400 points) are directly related to increases in the reliability of the entire map" (Walden, 1972, 93).

In contrast, Morrison suggested that,

> "twenty-five points are apparently too few for use in the isarithmic mapping technique, and for complex distributions, one should probably have a minimum of more than 49 known data points. There does not appear to be a need to exceed 100 known data points" (Morrison, 1971, 61).

Clearly, the size of data set required depends on the range of the area to be interpolated and the accuracy of

interpolation that is considered suitable. Morrison used inherently smooth surfaces, and therefore it might be assumed required less points adequately to define these surfaces. Additionally, since Morrison and Walden were interested in examining the relative effect of sample size on interpolation, their tolerances might be different to those used in this study.

It was decided initially that all the points necessary to define a specific sample type would be observed, thus varying sample sizes were generated. Subsequently, sample sizes of approximately 250 points would be randomly selected from the larger data sets and augmented by approximately 50 points on the border of the area of interest. This border sample would be common to all data sets, and provide a systematic decrease in error at the edges which would remove any 'edge-effect'. The size of the final samples (approximately 300 points) was considered a balance which would produce adequately differing results for the various terrain, sample type and interpolation combinations, while also providing a large enough sample to keep computation times operationally reasonable.

3.4 DATA ACQUISITION

3.4.1 The System

The photogrammetric models were observed on the Officine Galileo Stereosimplex IIc/Wang 2200s photogrammetric

digitising system in the Geography Department of the University of Glasgow (Petrie and Adam, 1980). The system contained the following elements:

a. incremental linear encoders attached to an Officine Galileo Stereosimplex IIc;

b. a small console containing nixie-tube x,y,z, co-ordinate display and the interface electronics;

c. a Wang 2200S desk-top computer with an 8K byte (8 bit word) memory;

d. a Baydel cassette drive to record data in ECMA 34 (ISO-3047) standard format;

e. a GTCO digitising tablet.

The Stereosimplex IIc is standard, with clamps in all three axes and slow motion screws in x and y. The encoders have a resolution of 5μm and give a constantly up-dated output to the nixie-tube display (5 decades and a binade), which is only transfered to the Wang on depression of the foot-switch. The standard display units also include preset and reset buttons and thumbwheel switches. The plotter could therefore be

'locked' in x,y, and z and the readings transferred to the thumbwheels. This allowed the digitising system to be switched off without loss of co-ordinates and origin, as the co-ordinates could be reset by depressing the 'reset' button while the plotter was locked at a fixed point.

Additionally, the cartographic digitising tablet could be directly connected to the Baydel unit, although in practice this was done through a VDU terminal which allowed data to be flagged as they were observed. No software was therefore used, with the baud rate of the digitising tablet controlling the flow of data in stream mode.

Once the ECMA 34 digital data cassette tape contained sufficient data, it was transferred onto 12.7mm, 9 track, IBM-compatible tape using the data transfer satellite available in the University of Glasgow Computing Service.

3.4.2 Method Of Data Acquisition

The general procedure for data acquisition was as follows. For each data surface, the model was established in the conventional photogrammetric manner (see American Society of Photogrammetry, 1965). However, since relative accuracy was of prime importance, great care was taken with the inner and relative orientations. Having established a good orientation, sampling was commenced. The various data sets (BREAKLINES, CONTOURS, GRID, RIVERS and SCATTER) were observed simultaneously (and flagged for later separation) in order to distribute any error,

a. due to the machine drifting out of orientation over the period of digitising,

b. due to strain on the observer while observing certain samples or at certain times of digitising. It was found that there was a human limit (about 6 hours for the author) which could be spent monotonously observing a grid without a greatly increased frequency of errors being made. In addition, an observer is fresher in the first few hours of observation and this might affect the resultant accuracy of some data sets.

While digitising, point mode encoding was chosen in order to:

a. allow careful selection of linear points in optimum locations without the urgency implicit in stream mode;

b. prevent any loss of data due to the inability of the system, given the existing small buffer sizes, to cope with stream digitising;

c. overcome some of the problems inherent in the then stream digitising implementation in the system.

Several features occurring while digitising are worthy of note. A continuous point count was augmented and recorded with the co-ordinates such that, at any stage, the total number of

points digitised within the model and within a specific sample could be assessed. Where linear data were observed, a point well outside the area of interest was chosen as a terminator to aid later post-observational processing. All the visible river systems were sampled at points of morphological significance and on both banks where possible. The major breaks of slope were sampled on points of morphological significance. The grid was observed by scanning in y and, within that, up and down in x, utilising clamps and slow-motion screws to ensure precise positioning.

The contour digitising incurred the most problems. Since it was impracticable to use stream digitising on the plotter, and since it was felt that a greater accuracy could be obtained by point digitising, the FORV contours were initially digitised in point mode, although a registered contour plot was also directly drawn on the attatched plotter table. The plot was subsequently digitised manually on the tablet digitiser. Since the original and digitised plots were indistinguishable, only the latter system was used for LIAN and INCH, as it was faster and resulted in smoother contours than the photogrammetrically point-digitised contours. The contours were digitised following MacLeod's rule that,

> "the (higher) area lies to the right of the line as digitised" (Keir, 1976, 13).

3.4.3 Post-observational Processing

A certain amount of post-observational data processing was required to convert raw observations into a flexible data

source. All data were transformed from digitiser co-ordinates into local terrain co-ordinates. Gridded data were additionally reformatted from x,y,z format to a sequential, compact store of z co-ordinates only. Any grid subset, and indeed grid-derived point sample could be generated from this format using the program 'REFORM' (see A1.4.18). Linear data were trimmed to the border and any polygons (e.g. contours) were closed. Finally, all data sets were re-sampled to produce similarly sized data sets (approximately 250 points), and sufficient border points were added to the data sets to permit interpolation without significant edge-effect distortion.

Table 3.2 and Figure 3.4 summarise all the data sets.

3.4.4 Accuracy

In order to understand fully any subsequent interpolation errors and to relate the data sets utilised within this thesis to 'reality', the exact nature of the accuracy of the data must be estimated. This may be considered in terms of photogrammetric and cartographic data capture errors.

3.4.4.1 Photogrammetric Digitising Error

Photogrammetric errors may be summarised in terms of instrumental, imagery, control and observational errors. It is generally considered that the planimetric instrumental accuracy of a Stereosimplex IIc is in the range 0.02% to 0.08%. These figures only represent the accuracy in stationary mode and are therefore unrealistic for dynamic mode, which Marckwardt (1978) suggests is the main reason for

photogrammetric plotting not meeting a required accuracy. Therefore, within this research, the accuracy of the contours would be expected to be slightly less than that of the other data sets.

Strain states that:

> "the relative accuracy of spot elevations and the running contours is a subject of considerable controversy. ... a summary of reports I have seen and studies I have been directly involved in indicate a relative factor of between 1.67 and 2 ... it seems reasonable to accept the 2x factor,..." (Strain, 1972).

In addition, non-instrumental error sources must be considered. These are:

a. ground control errors - both poor distribution of control and errors of position;

b. photogrammetric materials - movement of the emulsion base material stability, processing errors;

c. image collection errors - refraction, non-flatness of film, lens and residual distortion;

d. observational errors - control point mis-identification, scale reading errors, failure to completely remove parallax, backlash in the instrument.

The total effect of photomaterials and image collection

errors is generally considered to be approximately 10μm (Marckwardt, 1978). However, these errors and, more importantly, the ground control errors, will only affect the absolute accuracy of any co-ordinates obtained. Since the photogrammetric sampling was to derive data sets in local co-ordinates and not to exactly model a portion of the Earth's surface in relation to all other parts, these three error sources were ignored. Marckwardt (1978) stressed the importance of operator experience, or observational error, and this is of major significance in this study. However it is extremely difficult to quantify these errors; great care was taken to minimise them.

Additionally, the impact of the digitiser must be considered. While resolutions of 0.005mm are quite common on photogrammetric digitisers, Petrie (1970) suggests a more realistic figure of +/-(0.02-0.03)mm as the standard deviation obtainable in the model area. This would be slightly lower in dynamic mode.

In order to quantify all the observational, digitiser and plotter errors, three terrain points were re-observed ten times during observation. The resultant standard deviations obtained in model co-ordinates were 0.043mm in x, 0.031mm in y and 0.013mm in z. This is equivalent to deviations in terrain co-ordinates of 0.215m, 0.155m and 0.065m for FORV and 0.645m, 0.462m and 0.196m for LIAN and INCH. In order to overcome the problems of additional dynamic mode errors, sampling for FORV contours was attempted in stationary mode. When the contours were digitally plotted at final scale, it became apparent that

this was unsatisfactory, as the contours were too angular, although there was good correlation between the points and the cartographically digitised photogrammetric contours. Subsequent contours were only digitised cartographically from the photogrammetric plot. This combination involves similar accuracies of procedure as,

> "the accuracy which will be achieved in stereoplotting machines during the dynamic mode of plotting will be rather lower (than in stationary mode), but should be at least as good as that of any cartographic digitiser" (Petrie, 1970, 7).

3.4.4.2 Cartographic Digitising Error

Cartographic error, like photogrammetric error, may be sub-divided into instrumental and non-instrumental errors. Instrumental error varies with instrument resolution. Rogers and Dawson (1979) stated that current resolutions of cartographic digitisers vary from 0.0025-0.25mm. Petrie (1970) suggested that a standard deviation of approximately +/-0.1mm is generally attainable. The digitiser used for digitising contours was a Summagraphics ID tablet with resolution 0.13mm and an accuracy of 0.1mm.

In addition to digitiser error, errors caused by line width, paper stretch and handshake should be considered. A line width of 0.1mm on a low distortion plastic base was used in an attempt to make the first two errors negligible. As with photogrammetric digitising, observational errors represent a source of error which is difficult to quantify as a constant; however, several points (grid nodes) were re-observed in order

to obtain an estimate for the standard deviation of digitising, and thus an estimation of the observational and digitiser errors. The resultant standard deviations obtained were 0.09mm in x and 0.12mm in y. This is equivalent to deviations in terrain co-ordinates of 0.450m and 0.600m for FORV and 0.675m and 0.900m for LIAN and INCH.

Overall, it is felt that this accuracy is comparable with most of the disciplines in the earth sciences, mapping, and engineering. Connelly stated that

> "terrain is the most basic geographic feature" and that "extreme accuracy is seldom demanded in regional geographic mapping. When working in areal studies such as demography, climatography or ecology, geographers prefer to consider only the large scale regional features, thus the accuracy required of terrain height data is somewhat higher than that required of other geographic data" Connelly (1971, 129 and 101).

It has already been shown that the standard deviation of the data is comparable with that expected in present-day topographic mapping. In other earth sciences, accuracies of +/-1m are generally more than adequate. For engineering purposes, Phoneko (1971) suggested that an accuracy of spot heights in detailed design of 0.1m-0.2m is adequate. Additionally it is considered that photogrammetric height accuracy is

> "sufficient for all civil engineering projects" (Blaschke, 1969, 4).

Indeed, the data collected in this thesis would even pass the most stringent 1:500 scale specifications in the MOSS ground model system (Craine and Heatherington, 1977).

It must therefore be concluded that the data are comparable in terms of accuracy to that required within many other

disciplines, and, at least in terms of accuracy, are thus suitable samples for this examination.

3.4.5 Characteristics Of The Observed Data Sets

In the previous discussion, the derived data sets have been considered in relation to the sampling method of the observed surface. Discussion, while uniquely defining each data set within the context of the originally measured photogrammetric model topography, has not primarily attempted to define each data set in relation to each other data set. It is necessary to utilise objective repeatable measures to describe the data sets in order to:

a. be able to describe the data sets without relation to the enclosing shape or dimensions of the area.

b. be able to compare the data utilised within this research (and thus the results of using such data) to that utilised elsewhere and in other disciplines.

c. precisely describe the data sets for the future statistical analysis of error distribution and causation (see Chapter 6).

Unwin (1981) suggests that, the maximum amount of information can be extracted from point maps by utilising some method of pattern descriptor.

"Pattern is that characterictic of a spatial

> arrangement given by the spacing of individuals in relation to one another. It is, therefore, not related to the shape of the study area in which it is found and can be contrasted with the property of dispersion, which is the spacing of objects in relation to an enclosing shape. The density is the property of dispersal relative to an area but is independent of the area shape or the dispersion of the objects within it", (Unwin, 1981, 37).

Unwin suggests that three distinct approaches may be utilised to solve this problem.

3.4.5.1 Measures Based On Density

The simplest measure which may be adopted is crude density (the number of points per unit area). Since similarly-sized data sets of varying pattern which occupy the same area produce identical crude density values, it must be considered to give very little information about dispersion and pattern.

An alternative is to utilise quadrat analysis, and to generate a variety of statistics from a distribution of the number of data points falling into a series of identically-shaped and equi-sized sub-areas - or quadrats - laid across the data area. This produces a picture of density variation across the area and is generally described by Unwin (1981), Taylor (1977) and Ripley (1981). A computer program (QA - see A1.4.17) was written to generate some of the quadrat analysis statistics mentioned in the literature, and the results are summarised in Table 3.3a/b. The program is based on the more common 'quadrat censusing' - the area of interest being exhaustively partitioned into a grid of contiguous quadrats - every point being counted once only. Two cases are considered in the tables:

Table 3.3 Summary of quadrat censusing statistics.

DATA	MEAN	VARIANCE	RATIO	ICS	ICF	X^*	IP
IB	3.380	7.376	2.182	1.182	2.859	4.562	1.350
IC	2.900	2.750	0.948	-0.052	-55.877	2.848	0.982
IG	3.240	2.193	0.677	-0.323	-10.024	2.917	0.900
IR	2.940	9.870	3.357	2.357	1.247	5.297	1.802
IS	2.900	3.053	1.053	0.053	55.139	2.953	1.018
LB	3.050	12.573	4.122	3.122	0.977	6.172	2.024
LC	3.160	5.425	1.717	0.717	4.409	3.877	1.227
LG	3.040	4.634	1.524	0.524	5.799	3.564	1.172
LR	3.220	15.721	4.882	3.882	0.829	7.102	2.206
LS	2.940	4.912	1.671	0.671	4.383	3.611	1.228
FB	2.500	5.444	2.178	1.178	2.123	3.678	1.471
FC	2.500	3.343	1.337	0.337	7.410	2.837	1.135
FG	2.500	1.222	0.489	-0.511	-4.891	1.989	0.796
FS	2.500	4.818	1.927	0.927	2.696	3.427	1.371

a) using 100 cells.

DATA	MEAN	VARIANCE	RATIO	ICS	ICF	X^*	IP
IB	0.845	1.102	1.304	0.304	2.782	1.149	1.359
IC	0.725	0.630	0.870	-0.130	-5.559	0.595	0.820
IG	0.810	0.193	0.238	-0.762	-1.063	0.048	0.060
IR	0.735	1.659	2.258	1.258	0.585	1.993	2.711
IS	0.725	0.701	0.966	-0.034	-21.565	0.691	0.954
LB	0.763	1.360	1.784	0.784	0.973	1.547	2.028
LC	0.790	0.779	0.987	-0.013	-5.859	0.777	0.983
LG	0.760	1.192	1.568	0.568	1.338	1.328	1.748
LR	0.805	2.391	2.971	1.971	0.409	2.776	3.448
LS	0.735	0.925	1.259	0.259	2.838	0.994	1.352
FB	0.625	0.816	1.306	0.306	2.041	0.931	1.490
FC	0.625	0.631	1.010	0.010	65.628	0.635	1.015
FG	0.625	0.235	0.376	-0.624	-1.002	0.001	0.002
FS	0.625	0.803	1.285	0.285	2.191	0.910	1.456

b) using 400 cells.

a. where 100 quadrats overlayed the data area.

b. where 400 quadrats overlayed the data area.

From the resultant intensity distribution (points/quadrant), the mean may be generated. As mentioned previously, this conveys little information about the distribution pattern, but in this case only reflects the individual sample size. The variance of the intensity estimate depends on the spatial pattern of the data set. An even data distribution results in a low variance - thus for all the models, GRID data have a very low variance and RIVER and BREAKLINE data have the highest variance. A ratio (variance/mean) may be established to relate spatial pattern to density - this value increases with clustering.

Ripley (1981) discusses many other statistics which may be evolved from these basic statistics.

a. The ICS (Index of Clustering) is derived similarly to the ratio (ICS = variance/mean-1). If the population is clustered, the ICS will be large, whereas if the individuals are regularly spaced the index will be negative.

b. The ICF (Index of Cluster Frequency) measures the mean number of clusters per quadrant (ICF = mean/ICS).

c. $\underline{x}^*$ (Index of Mean Crowding, x^+= mean+ICS) represents the number of individuals sharing a quadrat with a typical individual in another cluster.

d. IP (Index of Patchiness, IP = x^*/mean = 1/ICF+1) represents the number of individuals sharing a quadrat with a typical individual in the same cluster.

At first sight this technique appears to generate a wealth of statistical descriptors of the pattern of the data sets, although these, and many other statistics suggested by other workers, are in fact similar concepts expressed slightly differently. Unwin (1981) outlines a few problems in using this technique. The frequency distribution strongly depends on the quadrat size, shape and origin; this may be apparent from considering the results in Tables 3.3a/b - varying one of the parameters (size) affects the final statistical results (see particularly the LIAN data sets). Another problem, particularly if statistical analysis is intended, concerns the two quite different ways of generating the frequency distribution. 'Quadrat censusing' was used in this research, but this produces differing statistics from the other common method - 'quadrat sampling'.

3.4.5.2 Measures Based On Distances

An alternative method of pattern description uses a spacing approach based on the distances between data points. This assumes a Euclidean geometry approach to deriving the

inter-point distances. A computer program (PTDIST - see Appendix 1.4.16) was written to compute several of the statistics within this group of techniques - the results are summarised in Table 3.4.

The simplest descriptor is the mean centre of the data set, generated simply by averaging all the x and y spatial co-ordinates for the data sets. By considering the third dimension (z-value) as a weighting factor, a weighted mean centre may be derived. As with ordinary statistics, the scatter of the data around the mean may be measured by the standard deviation of the data. Usually termed the standard distance and the weighted standard distance, these measure the spatial dispersion of the point pattern - the values increasing with dispersion. These may be computed separately (for x and y) or pooled. A disadvantage of these measures is that they are all scale-dependent and they would thus have to be normalised for inter-model comparisons. In addition, since within this research a relative distribution descriptor is of more interest than an absolute descriptor, the standard distances, particularly the separated (x and y) statistics, are of more interest. Superficially, these provide a low consistency descriptor.

Associated with this approach is the derivation of the mean of the distances between each point and its nearest, second, third- and higher-order neighbours, (see Table 3.4). This nearest neighbour approach, which has been considered for the first three orders of neighbours for the data sets within this research, is subject to many limitations, especially the

DATA SET	MEAN CENTRE		STANDARD DISTANCE			WEIGHTED MEAN CENTRE		STATISTICS STANDARD DISTANCES			MEAN NEIGHBOUR DISTANCE		
	X	Y	X	Y	POOLED	X	Y	X	Y	POOLED	1ST ORDER	2ND ORDER	3RD ORDER
1B	691.8	705.6	463.4	427.6	630.5	618.6	678.7	319.9	335.1	463.3	43.3	62.5	78.6
1C	734.8	704.0	440.3	442.4	624.2	672.2	691.5	329.9	334.6	469.9	55.1	80.8	97.5
1G	763.3	763.3	464.2	464.2	656.5	680.8	740.2	332.5	346.7	480.3	83.7	89.6	90.2
1R	712.6	748.8	451.6	467.8	650.2	651.3	741.9	324.6	346.4	474.7	35.2	50.6	68.5
1S	757.6	706.9	453.3	424.6	621.1	678.8	683.5	329.2	330.3	466.4	51.2	75.4	93.5
LB	336.1	542.8	243.4	270.1	363.5	356.1	594.2	178.8	230.9	292.0	24.7	31.8	40.3
LC	413.4	546.5	237.9	259.0	351.6	437.5	599.6	197.5	231.2	304.0	31.0	40.7	48.3
LG	384.4	450.0	224.2	261.1	344.2	403.9	505.7	184.2	206.1	276.4	44.0	47.4	52.6
LR	298.0	451.7	228.3	272.7	355.6	300.7	508.3	153.7	200.0	252.0	17.6	22.9	32.6
LS	357.2	511.1	226.2	276.0	356.8	377.2	567.2	183.2	224.7	289.9	26.3	38.4	48.6
FB	395.2	326.1	212.9	182.5	280.5	389.1	340.9	120.1	112.4	164.5	22.4	30.6	38.4
FC	409.7	354.4	201.7	195.8	281.1	403.2	370.1	123.7	118.5	171.4	26.7	34.5	41.8
FG	378.4	381.7	206.3	207.3	292.5	369.9	400.5	117.8	122.6	170.1	41.8	41.8	42.7
FS	396.2	345.1	211.7	192.8	286.3	390.2	360.6	121.5	116.8	168.5	23.4	31.6	39.1

Table 3.4 Summary of pattern statistics based on distance measures.

problems of lack of consideration of directions and edge-effects (distortion must occur at the border). Its best use is not as a 'snapshot' at a certain order of neighbour, but in establishing a trend. For example, the evenly-distributed GRID data have good consistency between the different orders, whereas the clustered RIVER data show a marked increase between each order.

This approach (of determining nearest neighbours) forms a basis for one of the most popular pattern statistics - Nearest Neighbour Analysis (NNA), see 3.2.1. It has the advantages that the statistic is dimensionless - describing the divergence of the actual point pattern from randomness (1.0) by using a simple ratio. Additionally, it is frequently used in the literature as it is easily understood and forms a unique, repeatable and consistent descriptor of pattern. A computer program (NNA) was developed to compute this statistic for the research data sets (see Table 3.2) and the results will be used in the later analysis of random-to-grid interpolation results (see Chapter 6).

3.4.5.3 Measures Based On Distance And Direction

Unwin also suggests a third category of methods which characterise pattern. While it may be agreed that the previous methods have their disadvantages which this final category solves - providing no information on relative location of individuals (quadrat analysis), and problems of edge-effect (distance measures) - it is believed (after experience with the Thiessen approach within 'DTC' - the self-developed

triangulation and contouring program) that the Thiessen approach suggested provides a less consistent and perhaps more subjective approach. Unwin suggests that,

> "Many other measures could be developed" (Unwin, 1981, 48),

and this has prompted the use of the (perhaps more traditional yet universal) NNA approach, as opposed to the more subjective and more varied Thiessen approach, whose use might be specific within this research.

CHAPTER 4

SURFACE CHARACTERISTICS AND ROUGHNESS

4 SURFACE CHARACTERISTICS AND ROUGHNESS

4.1 INTRODUCTION

> "In a general sense, roughness refers to the irregularity of topographic (or other) surfaces" (Mark, 1975b, 166).

The examination of roughness of certain surfaces is subsumed under the subject of geomorphometry, the science that attempts to define quantitatively the form of the Earth's surface. Though the purpose of this research is to consider results in the broader context of non-topographic surfaces, descriptors will be based upon the topographic ones for convenience - most of the principles of decription of surface roughness are identical. Evans (1972, 18) sub-divides this into Specific geomorphometry - "the measurement and analysis of specific landforms such as cirques, drumlins and stream channels", and General geomorphometry - "the measurement and analysis of those characteristics of landform which are applicable to any continuous rough surface." The former is of no particular concern within this thesis. In general, only relative global roughness descriptors for each data set are required, and these must be simple and brief where possible, thus excluding many of the geomorphometric descriptors which are often quite complex. However, often in the natural environment, surfaces are not homogeneous but heterogeneous in

their properties; thus a purely global descriptor is not adequate and some form of sub-global descriptor is required.

Reviews of the state-of-the-art are provided by Evans (1972) and Mark (1975b); no attempt will be made to repeat them. Instead, the main concepts of geomorphometry will be outlined and some of the techniques used to derive those concepts, which are used within this thesis to describe surface characteristics, will be outlined. The results of these methods will be evaluated in an attempt to discriminate between surface type and data sampling methods.

4.2 ELEMENTS OF GEOMORPHOMETRY

> "A single concise definition of surface roughness is impossible " (Hobson, 1967, 1), however a "review of the literature of general geomorphometry shows that all of its parameters may be defined in terms of altitude" (Evans, 1972, 22).

Several parameters exist, namely:

a. Grain - the largest significant wavelength of the topography;

b. Texture - the shortest significant topographic wavelength;

(These two parameters may be used to indicate the scale of horizontal variations in topography.)

c. Relief - a measure of the amplitude of topography; Evans (1972) proposed that standard deviation of altitudes

provides a stable measure of the vertical variability of the terrain;

d. Slope - Evans (1972, 36) stated that "slope is perhaps the most important aspect of the surface form, since surfaces are formed completely of slopes";

e. Aspect - the dispersion of slope orientation;

f. Regional convexity - the dissection and aeration of topography. Perhaps the most popular coefficient of dissection is the Hypsometric Integral, first derived by Strahler to estimate sub-surface volumes of drainage basins. However, skewness of altitude alternatively provides a good indicator (Evans, 1972).

4.3 OUTLINE OF TECHNIQUES

The techniques used by geomorphometrists to derive the parameters of surface roughness are quite varied, although four general techniques may be isolated. These are the generation of autocorrelation and correlation functions; the computation of surface areas and dispersion vectors; the fitting of pointwise polynomials and the establishment of

their derivatives and the use of some form of spectral analysis.

4.3.1 Autocorrelation And Correlation

Thomas asserted that

> "geometry of surface profiles can be specified completely by two parameters, the RMS roughness and the correlation length" (Thomas, 1975, 205).

RMS roughness is simply the standard deviation of the height frequency distribution. Alternatively, autocorrelation gives a measure of the grain and texture of the surface. The concept of correlation length is based upon the extent of near-monotonic decline of autocorrelation with increasing lag.

Autocorrelation involves correlating one variable (say u) with itself, as opposed to correlation where two variables (u and v) are correlated with each other. The variable may be one- or two-dimensional and the resultant plot of this autocorrelation function is known as a correlogram. The autocorrelation 'r' between u and itself is a measure with the following general properties:

a. $-1 \leqslant r \leqslant 1$;

b. r is positive if the values of u being correlated are associated positively, and negative if they are associated inversely;

c. The absolute value of r is 0 if the values being

correlated are not associated (uncorrelated), 1 if they are perfectly associated, and between 0 and 1 for intermediate cases.

There are several competing definitions used to define when the autocorrelation function has declined to an insignificant level. Whitehouse and Archard (1970) use 1/e of the initial value (+1.0) as the limit. Thomas (1975) uses 1/10 of the initial value. Drewry (1975) defines the limit at a given confidence level - usually 95%.

The actual autocorrelation function is generally based on the Pearsonian product-moment correlation coefficient which may be developed to derive many forms of the autocorrelation descriptor. For example, Henley (1976) derived areal autocorrelation coefficients from irregularly spaced data. Additionally McDonald and Katz (1969) derived a polar autocorrelation function in describing the ocean floor. Other users of autocorrelation funtions include Beckman and Spizzichino (1963) and Switzer, Mohr and Heitman (1964).

The autocorrelation measure provides a descriptor which is flexible enough to apply to both random and gridded data. Areal autocorrelation functions, and the resultant correlograms, were thus examined for all the point data sets and two-dimensional correlograms were examined for gridded data as possible roughness descriptors.

4.3.2 Surface Area And Dispersion Vectors

The computation of a surface area descriptor is based on

the hypothesis that surface area increases with surface irregularity. Thus, the surface area descriptor is a measure of slope variation across the surface. The work of Hobson (1967 and 1972) has perhaps generated most interest in this form of descriptor, although Turner and Miles (1968) and Hammond (1958) also considered this technique.

Hobson derived three parameters of terrain roughness,

a. an estimate of actual surface area compared with the corresponding planar areas;

b. an estimate of the 'bump' or elevation frequency distribution;

c. a comparison of the distribution and orientation of approximated linear surfaces.

The first two estimators are derived from a regular grid of heights. The last estimator is derived from irregular data which are in turn triangulated.

Surface area parameters may be derived by dividing the data area into cells, and comparing the surface area of each cell with its planar area. This parameter may be equally successfully derived for both point-based and gridded data samples. Since the surface area of each cell may never be less than the area of the planar cell, the minimum value of the descriptor must always be 1.0 (when the cell is flat). The

descriptor increases with cell steepness. On examining the distribution of the descriptors over the data area, the four moment-based statistics (mean, standard deviation, skewness and kurtosis) may be derived to enable a consideration of the irregularities of the data surface.

The four surface area statistics may be used to describe four different concepts. The mean is a measure of the average steepness of the terrain, thus identical values of the mean could indicate that the surface could be on a constant linear slope, or possibly range from very flat to very steep. Standard deviation is therefore used as a measure of the dispersion of the surface area. A low value suggests slope homogeneity, while a higher value suggests a more varied surface. Finally, skewness and kurtosis may be used to give more information on the shape of the distribution. Skewness describes the amount of symmetry or asymmmetry of the distribution, together with the direction of any asymmetry. Kurtosis describes the 'peakedness' of the distribution and thus the even-ness of the distribution. Its use is similar to standard deviation in that in principle it helps to describe the homogeneity of the distribution. To compare the degree of deviation from a normal distribution, skewness is without doubt the most important single statistic, followed by kurtosis (Evans, 1979).

By the very nature of the individual cell surface area descriptor, the frequency distribution is truncated at 1.0. Additionally, after preliminary examination, it was always found to be strongly positively skewed. There was therefore a

possible danger that the power of the statistics could be diminished. It has been established in similar work that non-zero skewness produces a bias to positive kurtosis and that they are positively correlated (Evans, 1979, 47). Additionally, in a skewed distribution, the mean is not so typical of the distribution as the median or, more importantly, the mode. There are various possibilities for over-coming such limitations. Evans suggests that quantile-based rather than moment-based statistics are worth investigating as descriptors. Alternatively, a transformation of the distribution values may give adequate results, reducing the skewness and therefore the effect on the mean and kurtosis. These aspects were considered important enough to merit further investigation. Median and quantile-associated statistics were also established, and the effect of various transformations on the frequency distribution investigated. These will be discussed later.

As an alternative to the surface area approach, dispersion vectors may be generated by similarly sub-dividing the surface into linear surface cells - usually triangles. Normals to these linear surfaces are represented by unit vectors. Thus mean vector orientation, vector strength and vector dispersion may be computed using methods defined by Fisher (1953) and described by Watson (1957). The total vector strength (R1) is obtained by using direction cosines and indicates the length of the resultant sum of the unit vectors. For the purposes of matrix comparison, this must be standardised, producing the standardised vector strength (R), which ranges from 0 (no preferred orientation) to 1 (identical orientation). Fisher's

dispersion factor (K) indicates the variability or spread of the unit vectors, and assumes low values for highly dispersed distributions and extremely high values for low dispersions.

Overall, 'smooth' areas are characterised by high vector strengths and low dispersions (high K values) and thus flat and tilted surfaces are not discriminated. Alternatively, 'rough' areas generate low vector strengths and high vector dispersion (low K values).

Turner and Miles (1968) also computed a terrain variability measure 'V' to enable an absolute roughness comparison between areas using the formula,

V = (range of elevation) / log (K) .

Since values of K can range from 1 to extremely large numbers, the variability factor can range from 0, for completely flat surfaces, to at least the range in elevations for extremely rough areas. Some of the results published by Turner and Miles are presented in Table 4.1 in their metric equivalent.

Dispersion vector analysis suffers from a similar statistical problem to surface area analysis in that the resultant frequency distributions are inherently skewed. Unlike surface area analysis, however, the vector dispersion distribution was consistently, modestly, negatively skewed. Similar techniques - transformations, medians, and quantile-based statistics - were considered in order to improve the resultant statistics. However, it was not felt that they were so appropriate in the context of vector analysis, since it produces absolute measures.

TERRAIN TYPE	ELEVATION RANGE (m)	ROUGHNESS (K)	VARIABILITY V (m)
ESCARPMENT	96	85	49.8
ESCARPMENT	85	82	44.4
HILLS	104	397	39.9
HILLS	84	220	36.0
HILLS	65	158	29.4
DRUMLINS	56	222	23.8
KARST PLAIN	51	723	17.7
KARST PLAIN	52	1329	16.6
PLATEAU	25	1759	7.6
MORAINE & DUNES	16	2892	4.7
GROUND MORAINE	18	24664	4.2
OUTWASH PLAIN	9	18024	2.0
GROUND MORAINE	8	30044	1.8

Table 4.1 Variability associated with particular terrain types (after Turner and Miles, 1968).

It was felt that an estimator based on surface area was easily comprehensible and resulted in a simple descriptor of surface slope. Similarly, some measure of the directional distribution of slope or aspect over the whole surface could be derived by examining the distribution and orientation of approximated planar surfaces. These techniques were therefore investigated further.

4.3.3 Pointwise Polynomials

More recently, Evans (1979) has established an 'Integrated system of terrain analysis' to generate a comprehensive evaluation of all the parameters of geomorphometry. The basis of this technique is the initial determination, from a regular DEM grid, of a series of pointwise quadratic polynomials centred on each of the internal grid nodes (altitudes) and covering a 3x3 node sub-grid. The first derivative of these polynomials is used to generate aspect and gradient directly and the second derivative to derive plan and profile convexities. Means, standard deviations, skewness and kurtosis are subsequently derived for the five variables (altitude, gradient, aspect, plan and profile convexity) and are displayed with the aid of scattergrams, histograms and line plots.

As with surface area analysis, this technique may produce consistently skewed distributions with an adverse effect on the resultant statistics. Evans (1979), while avoiding the use of quantile-based statistics, has undertaken considerable study concerning the effect on the statistics of utilising

various transformations. For altitude and gradient statistics, he concluded that,

> "even within one type of data, different transformations were required in different areas" (Evans, 1979, 40).

Such transformation of aspect distributions is not applicable, since moment-based descriptors are not appropriate for circular-scale measurements. As regards convexity, Evans considered it disadvantageous to apply quantile-based statistics, since they might hide interesting differences in the extreme tails. In this respect, the non-transformned moment-based statistics are generally satisfactory as long as the skewness/kurtosis bias is appreciated.

This method, in its present implementation, depends on the existence of a regular grid of heights. Thus its direct application is very much limited within this thesis. However, its use raises a problem of some importance - do interpolated grids produce similar values to those from directly observed data? Evans' integrated system of terrain analysis was thus utilised to analyse interpolated and observed grids.

4.3.4 Spectral Analysis

Spectral analysis is perhaps the most established technique, involving the fitting of a Fourier series (see 2.4.1.2.2.4) to the data to establish the texture and grain of the data surface. Rozema (1969) derived a Power Spectral Density function (PSD) to describe lunar surface roughness, similarly Rayner and McCalden (1972), Evans and Bain (1974a and 1974b) and Ayeni (1976) utilised the method for terrain

anlysis. In particular, Ayeni computed a combination of a linear filter and a double Fourier function and hence generated the 'Harmonic Vector Magnitude' (HVM). The HVM is defined as the square root of the sums of the squares of the coefficients.

Some criticism of the use of Spectral Analysis has been made. Evans has claimed that

> "such a (spectral analytic) model is natural for many time series generated by super-imposed periodic processes (especially in electrical engineering) and for the study of ocean waves, but there is no process justification for its application to arbitrary terrain profiles" (Evans, 1979, 90).

Additionally, Evans has argued that

> "the same information is implicit in mean gradient and vertical dimension" (Evans, 1979, 96).

Spectral analysis in its basic form was therefore not evaluated. However, a closely related technique, scale variance, was considered in more depth.

4.3.5 Scale-Variance

The estimators have so far been concerned with global measures, which inherently result in filtering. Surfaces are generally heterogeneous and not homogeneous and thus some form of sub-global measure is equally necessary. Indeed a sub-global measure would perhaps remove some of the concern that Hobson and other workers have shown over the single-parameter surface description.

Variability at different scales, and thus the heterogeneous nature of terrain, may be studied using nested analyses of variance (Moellering and Tobler, 1972). These result in the

establishment of a grain/texture descriptor. The Moellering and Tobler technique is an iterative procedure involving the use of analysis of variance techniques to examine scale effects in data structures. The implementation is restricted to a regular gridded distribution, since irregular distributions would require some *a priori* regionalisation to take place, with a scale-variable loss of data at the margins of the area.

The practical implementation of this technique is discussed in more detail in A1.4.20. In brief, at each level the sum of squares of the local deviations from the mean is computed and, using the degrees of freedom, the mean square and thus the scale variance component is established. This component gives a measure, when compared with the other components, of the magnitude of variance at respective levels and thus a measure of the heterogeneity of the surface. If the scale-variance component is similar at each level, or slightly decreasing, then a relatively homogeneous surface exists. If it varies irregularly, then a more heterogeneous surface exists.

The scale-variance component was considered to be essential to give a measure of the heterogeneity of the surfaces, although its implementation is limited to regular gridded input.

4.3.6 Conclusion

While there are five types of surface descriptor available, not all descriptors are satisfactory or are suitable for use

with all of the data sets that were used both within this thesis and in the 'real world'. The ideal descriptor must be capable of being evaluated both from regular gridded and also random point data. Where the descriptor may only be generated from a grid, the effects engendered by using an interpolated grid must be comprehensively assessed. Where both grid and point data may be used, some form of comparison must be possible of their relative performances.

4.4 EVALUATION OF DESCRIPTORS

In evaluating the descriptors, several pertinent points must be examined, namely:

a. the ability of the descriptor to distinguish uniquely the three data surfaces;

b. the ability of the descriptor to distinguish between the data sets of the different sampling methods - rivers, contours, grids, scattered, breaklines - both within each surface but also across the surfaces;

c. the ability of 'regular grid' descriptors to react similarly to observed and to interpolated grids. For this purpose, grids were interpolated using multiquadric analysis, the method that gave the most consistently accurate results for random-to-grid interpolation;

d. the consistency of the descriptors and their representation;

e. the quality of the statistically-based results, particularly with relation to surface area analysis.

4.4.1 Areal Autocorrelation Functions

The areal autocorrelation functions were evaluated using the program 'AUFN' (A1.4.1) which calls the subroutine 'ARAUT' as defined by Henley (1976). The program may compute MORAN or GEARY coefficients as outlined by Cliff and Ord (1973) although following tests, it was observed that the shape of the resultant correlograms were identical. Thus only the MORAN function was examined in relation to the various data sets.

In comparing correlograms, five features appear to be pertinent, namely

a. the initial slope of the curve;

b. the existence and location of a major shoulder on the curve;

c. the subsequent slope of the curve and the existence of any subsidiary shoulders;

d. the location of the intersection of the correlogram and the null correlation axis;

e. the location on the correlogram when 50% of the variance is accounted for, i.e. the correlation is 0.707.

4.4.1.1 Inter-area Comparisons

Examining the results for the three surfaces in general (Figures 4.1 4.2, 4.3 and with reference to Figure 3.3), there exists little difference between the LIAN and INCH data sets over the first 150m. This is not surprising since both areas are composed of macro-topographic features. In contrast, the micro-relief sand dunes of FORV contain much sharper, more irregular surface changes, thus resulting in a much more sudden decay, approximately three times the curve gradient of INCH and LIAN.

Similarly, FORV differs in its lack of shoulder in the correlogram. In all the LIAN data sets, a marked shoulder occurs at approximately 150m, due to the dimensions of the flat river basin in the South of the area and, since the hill slope rises very gently from this basin, data points (especially in the East-West direction) are liable to be highly correlated. The combination of these two features has the effect of high initial correlation, followed by a more

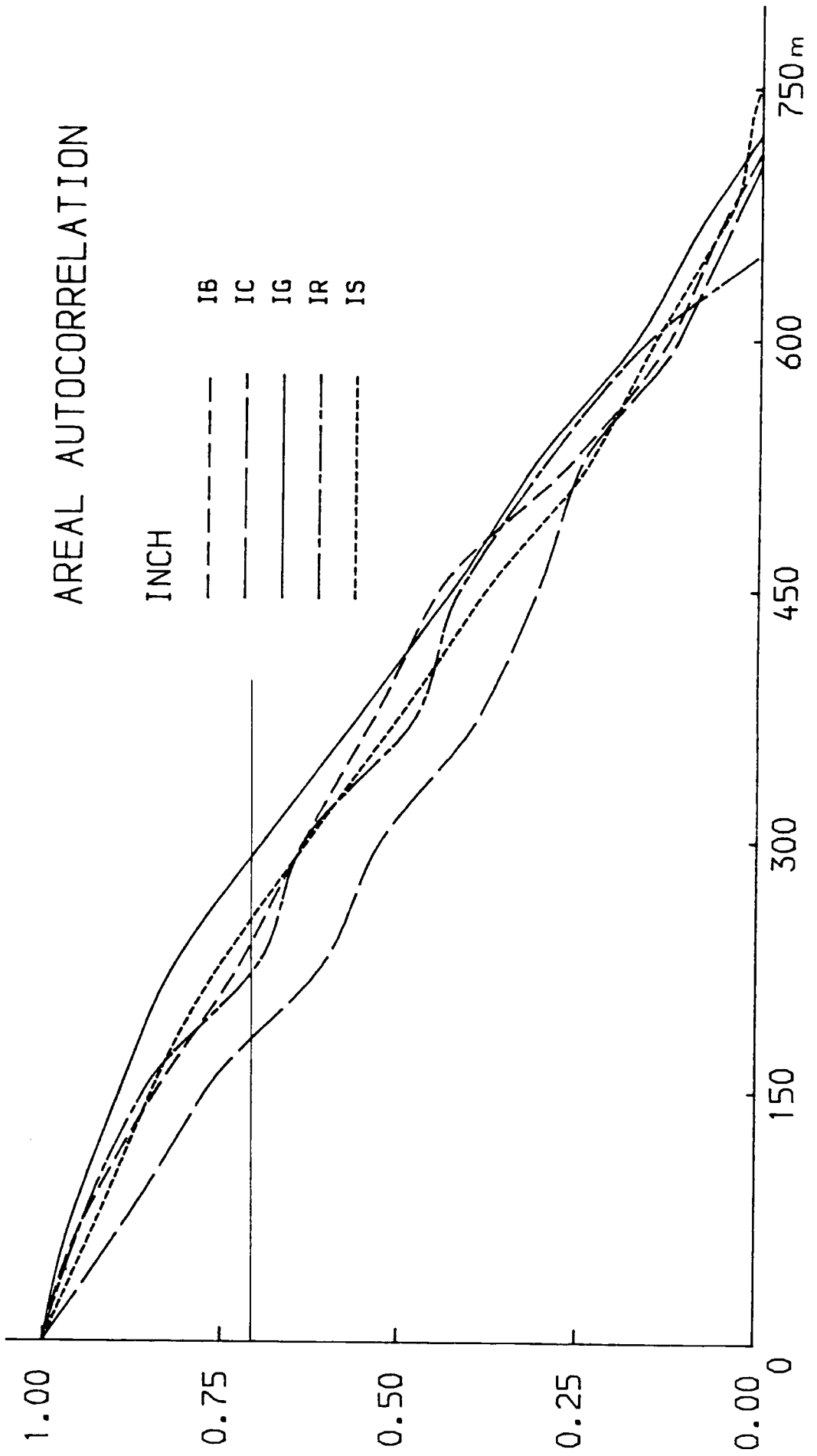


Figure 4.1 Areal autocorrelation - INCH

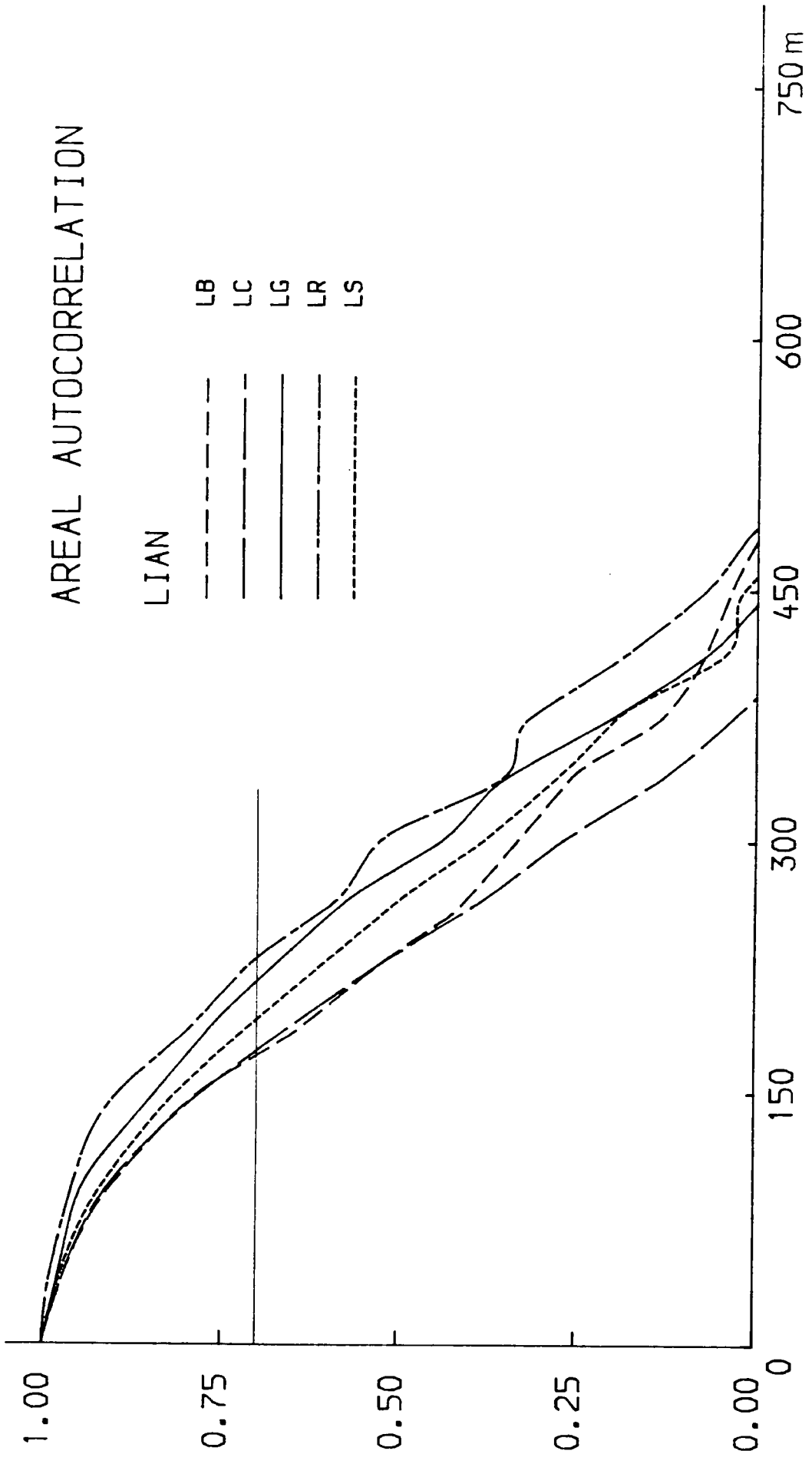


Figure 4.2 Areal autocorrelation - LIAN

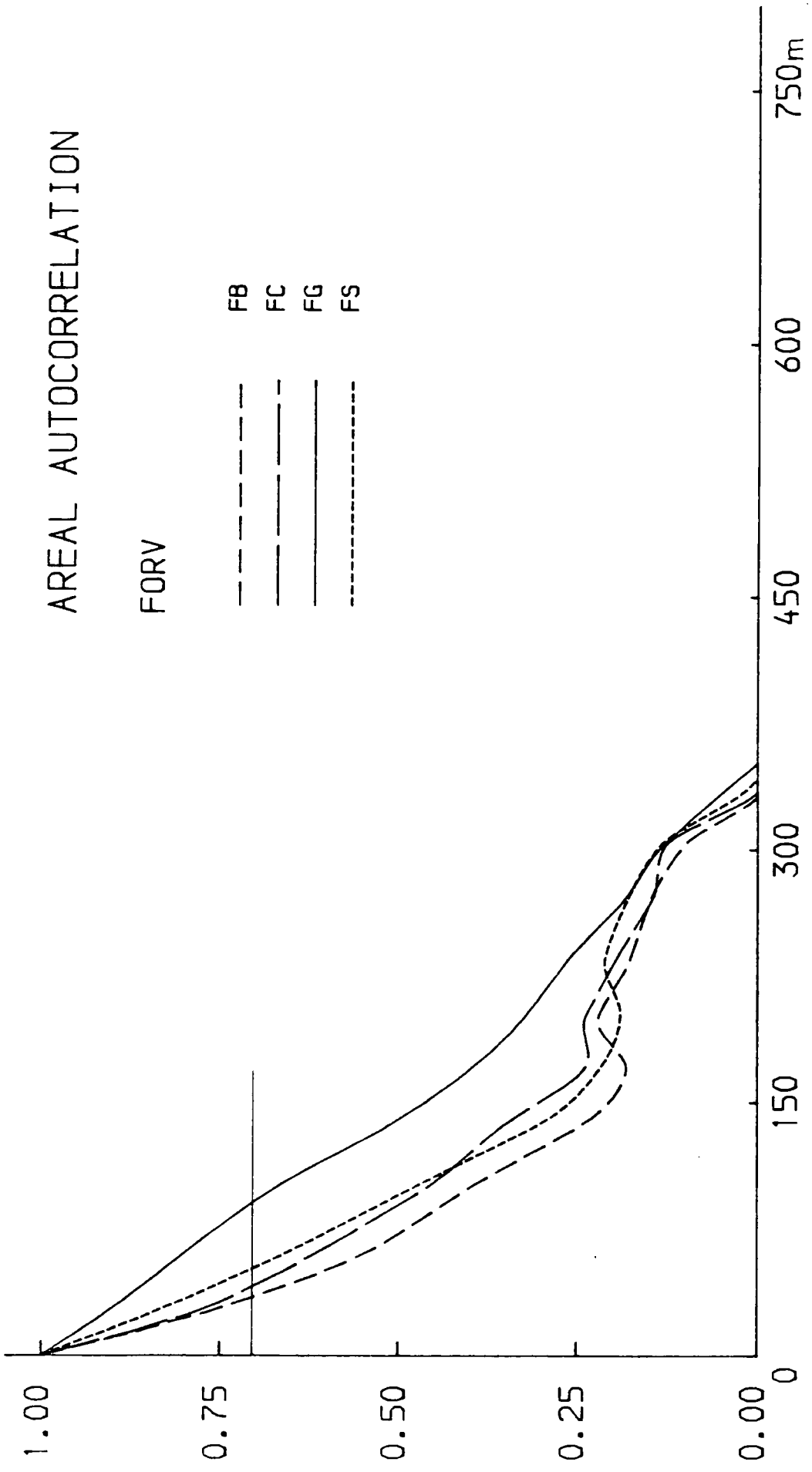


Figure 4.3 Areal autocorrelation - FORV

rapid decay, since at the longer distance basin points are correlated with hill-slope points. The INCH data sets have a poorly defined shoulder occurring at approximately 225m. Due to its poor definition, this is more difficult to explain. The INCH area is more dissected than the other data sets into well-defined units eg. river basin, river valley, scarp slope and other gently undulating hill slopes. Each has a well defined boundary, thus creating higher correlations within the boundary and marked decay across the boundaries.

FORV is also significantly different in that its correlogram shows a marked subsidiary shoulder at the 225-300m level. Consideration of Figure 3.3b shows that this is approximately the topographic wavelength in the East-West distance, the ridges and depressions being approximately 200m apart. LIAN and INCH have a smooth monotonic decline in the lower part of their correlograms, with only minor deflections of little importance.

Perhaps the most noticeable difference between the correlograms is where the curve crosses the 'x'-axis. The point at which null correlation exists is dramatically different for the three surfaces. FORV and LIAN are most similar, the shoulder occurring around 340m (FORV) and 450m (LIAN); in the INCH data set it occurs at approximately 700m. While this has no statistical importance, it does indicate that the decay in similarity in the case of INCH is far less than in the FORV and LIAN data (which should produce superior results if any interpolation is performed on an INCH subset). While this may be expected in the case of FORV, LIAN appears

to be a smooth, highly correlated surface. The 'poor' LIAN null correlation distance is, in fact, a result of using a global measure, since LIAN is in fact two separate surfaces with planar and linear sections. Two local measures reflecting this division might have resulted in separate, very high correlations.

Of a greater statistical importance is the point on the correlograms when 50% of the variance is accounted for. This is indicated by a short horizontal line at 0.707 on the correlation coefficient axes. While the FORV data sets are markedly different at this point, only a minor variation exists between LIAN and INCH data sets. This is a result of the lack of the shoulder in the INCH correlogram (as previously discussed), although it is exaggerated by the more rapid decay of the INCH-CONTOUR data set. This is emphasised by considering the mean decay distances to the 0.707 correlation value for each data area (FORV - 52m, LIAN - 196m, INCH - 236m).

4.4.1.2 Inter-data Set Type Comparisons

There appears to be a less consistent difference between data sets within each model than there is generally between different models.

Over the initial slope of the correlogram, all curves are relatively smooth and their paths do not cross those of other data sets. Table 4.2 summarises the order of autocorrelation in different parts of the curve. In general GRID points tend towards a higher level of autocorrelation than most other data

	INITIAL CURVE DECAY			DISTANCE TO 0.707 CORRELATION			DISTANCE TO NULL CORRELATION		
	INCH	LIAN	FORV	INCH	LIAN	FORV	INCH	LIAN	FORV
SLOWEST/FASTEST	GRID	RIVER	GRID	GRID	RIVER	GRID	SCATTER	RIVER	
					GRID			BREAKLINE	GRID
	RIVER	GRID							
	BREAKLINE			SCATTER			GRID		
		SCATTER	SCATTER	BREAKLINE	SCATTER	SCATTER	BREAKLINE	SCATTER	SCATTER
	SCATTER						CONTOUR	GRID	
		CONTOUR		RIVER					CONTOUR
								CONTOUR	BREAKLINE
			BREAKLINE		CONTOUR	CONTOUR			
	CONTOUR	BREAKLINE	CONTOUR	CONTOUR	BREAKLINE	BREAKLINE	RIVER		

Table 4.2 Within-model comparison of decay of areal autocorrelation for point data sets

sets, although for LIAN, RIVER data are more autocorrelated, and for INCH, the RIVER data are the second most autocorrelated. This is not surprising, since the points selected on rivers are usually quite close together and include only a small range of the total height range within a model. This is especially so in LIAN, where a sizeable number of RIVER points are along the valley floor. SCATTER data are generally the next most highly autocorrelated data set and this is almost certainly a result of the range of height values being greater than RIVER points, while being more evenly distributed across this range than, for example, BREAKLINE. BREAKLINE and CONTOUR generally produce similar initial curves, except in the case of INCH, where BREAKLINE data have a surprisingly high initial autocorrelation. The explanation for the generally low autocorrelation for BREAKLINE, especially over short distances, is undoubtedly as a result of the points being selected at the tops and bottoms of slopes. The nearest points will often be at the other end of the slope and therefore be considerably different. The inexplicably high autocorrelation for INCH-BREAKLINE may be a result of a denser distribution of points along fewer lines, thus the nearest point is not at the other end of the slope, but further along the ridge or valley. The low autocorrelation within the CONTOUR data set is more difficult to explain, but seems likely to be a result of the closest points being on different contour levels - except in LIAN, where a large number of points are in the river basin and have height values close enough to affect the function over the short distances.

The existence and location of shoulders on the functions

and the generalisation of the function vary between the data set types. GRID data give the smoothest autocorrelation function in each model. SCATTER data give the next smoothest curve, although in the FORV model this is nominally so. In addition, a slight 'toe' effect may be seen on both LIAN and INCH data sets where the function sharply changes direction as it nears a null correlation. In contrast, especially with the INCH model, the BREAKLINE, RIVER and CONTOUR data produce more 'wavy' functions, defining the characteristic wavelength of the particular topographic feature. In general, the location of the first shoulder of the functions appears to be a result of the model type rather than of the data set.

The strongest inter-data correlation occurs at the point on the correlograms at which 50% of the variance is accounted for. Excluding RIVER data, a strong positive correlation occurs between the ordered decay distances for each data set type. The unique circumstances concerning the localised and restricted nature of the LIAN-RIVER data set have already been discussed. This has the effect of making the associated correlogram decay less rapidly, and thus LIAN-RIVER is unrealistically high in the relevant section in Table 4 2. In general, the decay distance at which 50% of the variance is accounted for appears to be related to data type within any one model type. The decay distance where the value of the correlogram becomes zero appears to have least consistency. With the exception of CONTOUR generally having a shorter null correlation distance, no regular pattern appears to emerge (see Table 4.2).

4.4.1.3 Descriptor Consistency

Finally, the consistency of the descriptor must be examined. In order to measure this, the sizes of three of the data sets in each of the models were increased approximately two-fold (by re-sampling the original photogrammetrically-observed data) and the autocorrelation function re-computed. The new BREAKLINE, CONTOUR and GRID data sets produced the functions shown in Figure 4.4.

In all cases, the larger data sets produced a function which is of remarkably similar shape and runs near-parallel to the smaller data sets. These have values consistently smaller, with the exception of LIAN-GRID and INCH-CONTOUR. In particular the initial curve slopes (0.707 correlation points) and final null-correlation points are virtually identical, and slight changes are only discernible where minor subsidiary shoulders occur. Overall, the general patterns shown by each of the three surfaces are adhered to and the same types of data tend to be more distinctly similar over the first section of the curves; thus, for example, GRID data appear more distinct than BREAKLINE and CONTOUR data.

Thus, on the basis of these results, providing the number of data points within a data set is known, resultant autocorrelation functions from that data set can be reasonably matched to other data sets from that surface, and also, to a lesser extent, to similar data types in other data sets. Similarly, different data sets from different surfaces produce

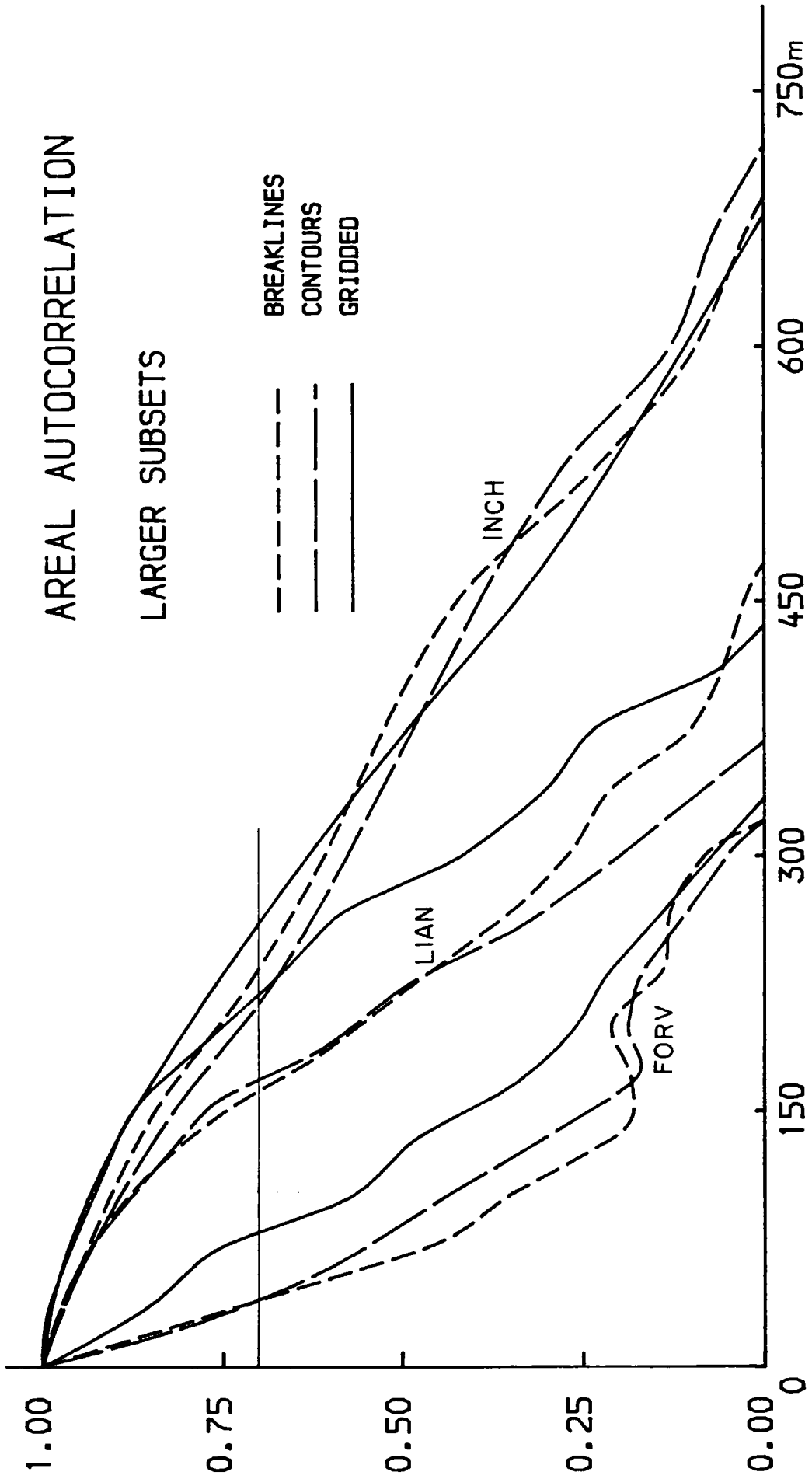


Figure 4.4 Areal autocorrelation - Larger data sheets

consistent results.

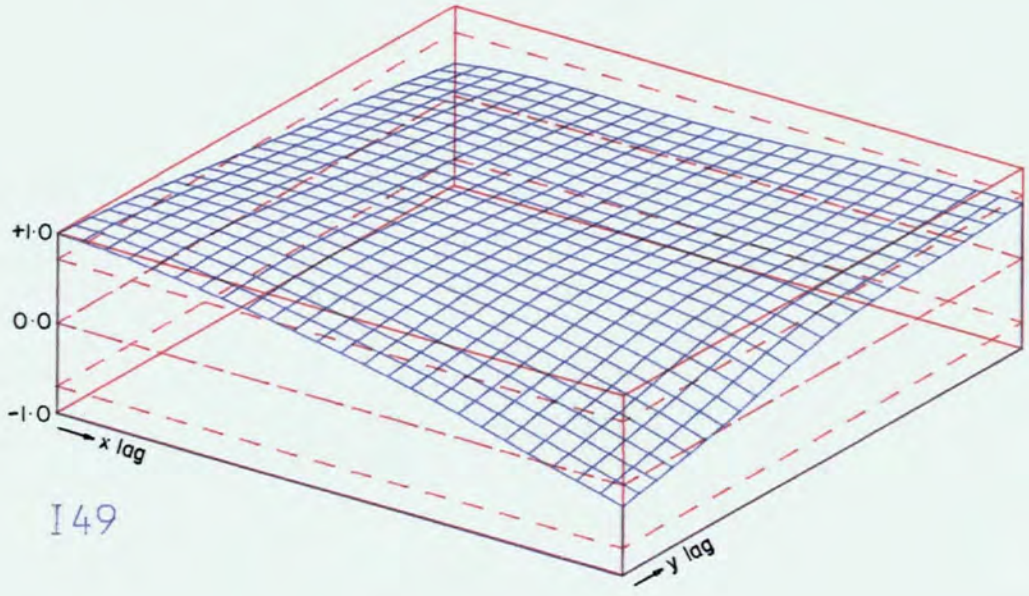
4.4.2 Two-dimensional Autocorrelation

One problem of computing a single areal autocorrelation coefficient is that it assumes that the data are isotropic. Unfortunately, this is not always necessarily true; however, in the absence of evidence to the contrary, it must be assumed. For the purposes of examining the autocorrelation of the data further, a program 'AUTOP' was developed by the author utilising the 'PERSYS' isometric block diagram plotting package (A1.4.2). The input to the program consisted of a regular grid; varying resolutions of observed grids and interpolated grids were input. The output from the program consisted of the two-dimensional isometric plot of correlations and the printed correlation matrix. The actual correlation function used was a modification of the Pearsonian product-moment correlation coefficient (Cox, 1979).

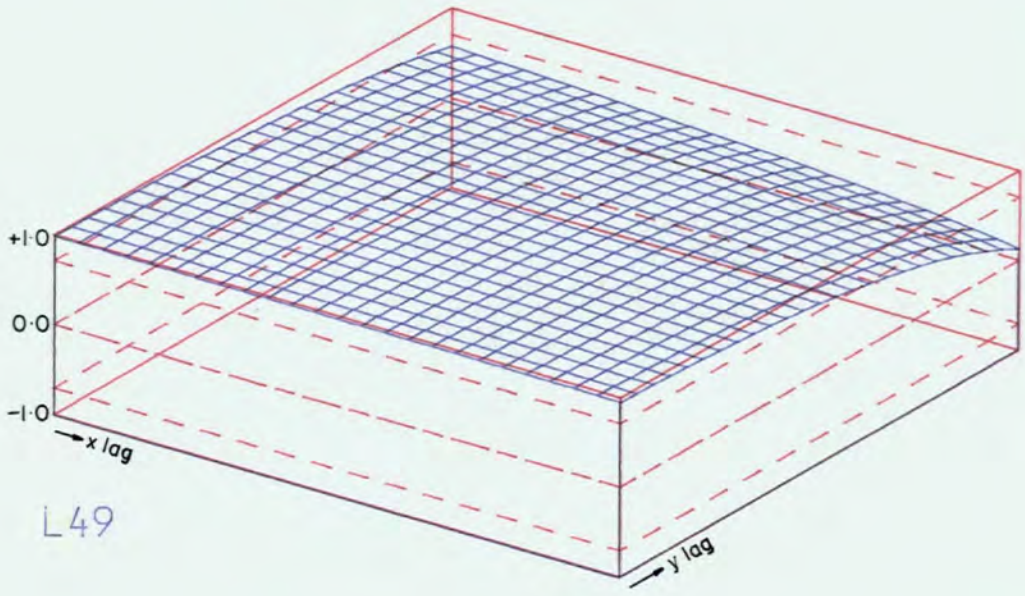
The isometric plots (Figure 4.5) are observed from the SE corner and thus the lines running bottom-left/top-right show the lag in 'y', and those running top-left/bottom-right show the lag in 'x'. The solid red lines illustrate the +1 and -1 limits of the coefficient, and the pecked lines 0 and limits +/-0.707. (where 50% of the variance is accounted for). The values where the function crosses +0.707 on lag 0 in both x and y are summarised with the corresponding AUFN distances for all the data sets in Table 4.3.

The table shows that a remarkable consistency exists within any data surface, irrespective of resolution. Thus the three

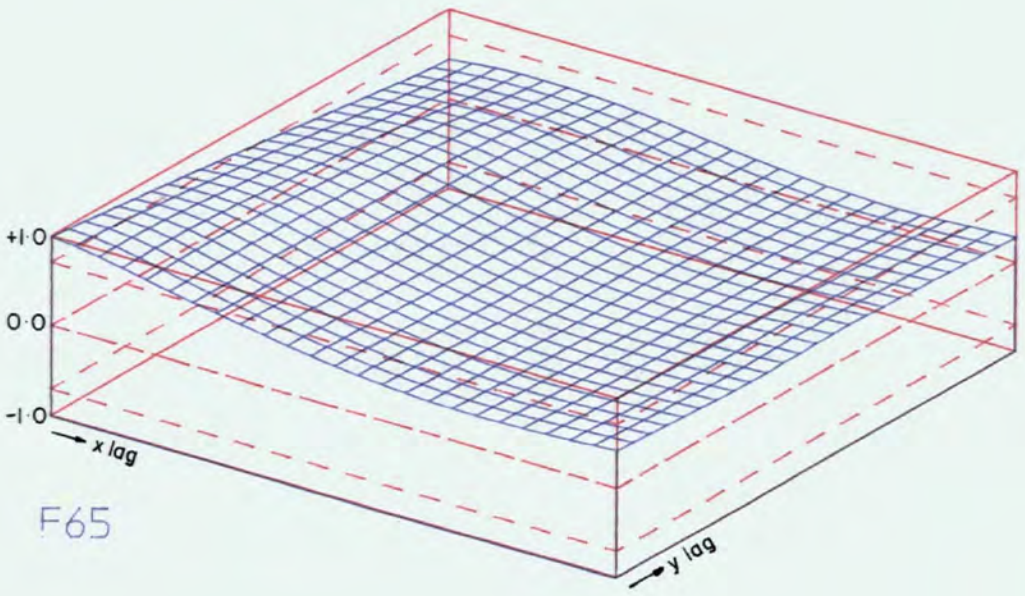
Figure 4.5 Two-dimensional correlograms



I49



L49



F65

DATA SET	LAG (cells)		DISTANCE (m)			AUFN (m)
	x	y	x	y	vector	
I49	9	11	270	330	426	
I25	4.5	4.7	270	280	389	
I13	2.3	3	280	360	456	
IB-MUL	9.3	12	279	360	455	237
IC-MUL	9.2	11	276	330	430	181
IG-MUL	9.5	12	285	360	459	293
IR-MUL	9.3	11	279	330	432	218
IS-MUL	9.2	11.8	276	354	449	251
L49	24 *	22	360	330	488	
L25	12 *	11	360	330	488	
L13	9 *	5	360	300	469	
LB-MUL	24 *	22.1	360	334	489	172
LC-MUL	24 *	21.2	360	318	480	177
LG-MUL	24 *	22.5	360	337	493	213
LR-MUL	24 *	24 *	360	360	509	227
LS-MUL	24 *	22.3	360	331	491	191
F65	6	13	60	130	143	
F33	3	6	60	120	134	
F17	1.5	3	60	120	134	
FB-MUL	7	13.1	70	131	148	32
FC-MUL	7	13.8	70	138	155	39
FG-MUL	7.4	16	74	160	176	87
FS-MUL	7.5	12	75	120	141	50

Table 4.3 Results of autocorrelation analyses
(* all lag values are above significant limit).

surfaces may be summarised by the respective plots for I49, L49, and F65. Minor differences are shown between the different data types and when the resultant vector is derived from the x and y distances and compared with the areal autocorrelation values at +0.707, a good correlation is shown to exist allowing for such small data sets (INCH - 0.91, LIAN - 0.91 and FORV - 0.47 - using Spearman's rank correlation).

INCH grids are shown to be reasonably isotropic with the decay reaching 0.707 at 270m lag in x and 330m lag in y. Beyond this, decay is rapid with the exception of the diagonal consisting of equal lag in x and y. This can be readily related to the actual INCH surface, as the general trend of the surface runs SW/NE following the line of the major dissection, the river valley. In addition, the more rapid decay along the major axes is caused by the existence of the same river valley, the other valleys transverse to that and the major landform change in the NE of the model. This suggests that, if interpolation is to take place in this surface, point search should preferentially occur in the SW/NE direction with lower weighting to other directions.

LIAN grids have a much higher level of autocorrelation than those from INCH or FORV, and greater than that suggested by the areal autocorrelation results. The E/W direction is confirmed as representing the major strike of the surface, as the autocorrelation is consistently higher in this direction. Conversely, the function decays most rapidly in the y (N/S) direction emphasising the dip of the surface. The AUTOP plot suggests that point searching for interpolation would best be

limited to a W/E direction, although a limited directional search in the N/S direction would also be acceptable.

FORV shows a more rapid decay in the W/E direction and a much slower decay in the N/S direction, which is far more rapid than for LIAN and INCH. Once again, this is easily interpreted from the terrain, where the general trend of the sand dunes runs N/S, hence the slower decay in this direction. Thus, for the purposes of ideal interpolation the point search should be concentrated in the N/S direction with low weighting in the W/E direction.

It has thus been shown that the two-dimensional autocorrelation function provides a descriptor which is very stable to both grid resolution and data type, and yet is very variable to surface type. This form of descriptor therefore provides an excellent indicator with which to control various aspects of interpolation, including both searching and weighting. As regards the within-model data type consistency, it is clear that a good rank correlation exists between the areal autocorrelation results and the two-dimensional (vector) results (Table 4.3). This could perhaps be improved by modifying the method for deriving the latter quantity. A more precise method would be to use the distance to the nearest inter-lag point where the value of the function is 0.707. Nevertheless, it does suggest that the interpolated grid still retains the basic character of the original point data set and that both autocorrelation techniques are mutually equivalent for the different forms of data required.

4.4.3 Surface Area Analysis

For the purposes of analysis of the surface area of the data sets, two programs were developed - 'SA' to deal with regular gridded data and 'DTSA' to deal with irregular point data. A full description of the programs occurs in A1.4.19 and A1.4.7. Both programs determine the amount of similarity between the surface units (triangles for DTSA, and square grid cells for SA) and assumed planar units (triangles or squares) using the parameters discussed earlier.

4.4.3.1 Regular Grids

A preliminary study of the histograms of surface area in a range of regular grid resolutions (Figure 4.6) and a range of surface models (Figure 4.7) shows that the distributions are strongly skewed. Some of the problems of utilising statistics derived from untransformed, truncated distributions with substantial skewness have been outlined earlier - means and standard deviations are distorted and kurtosis is biassed. The derivation of medians and inter-quartile ranges may aid the interpretative situation, although alternatively the moment-based statistics may be improved by transformation. It is probably unwise to apply quantile-based statistics in such a situation, as this would remove extremes which are an important aspect of the distribution.

The surface area statistics, like the other surface characteristics, were being utilised in two contexts;

a. The power of the method in providing a surface

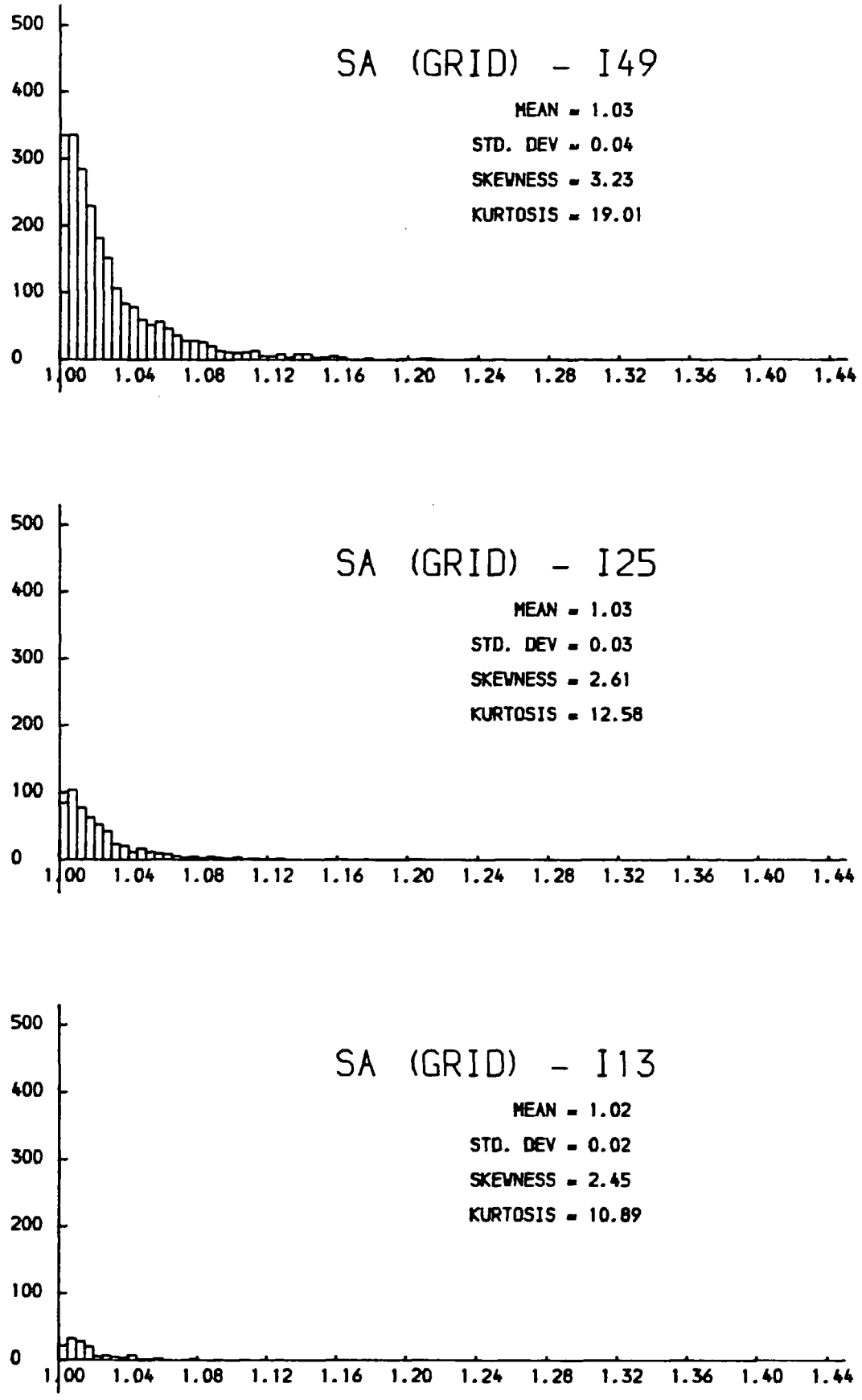


Figure 4.6 Surface area analysis distribution - range of grid resolutions

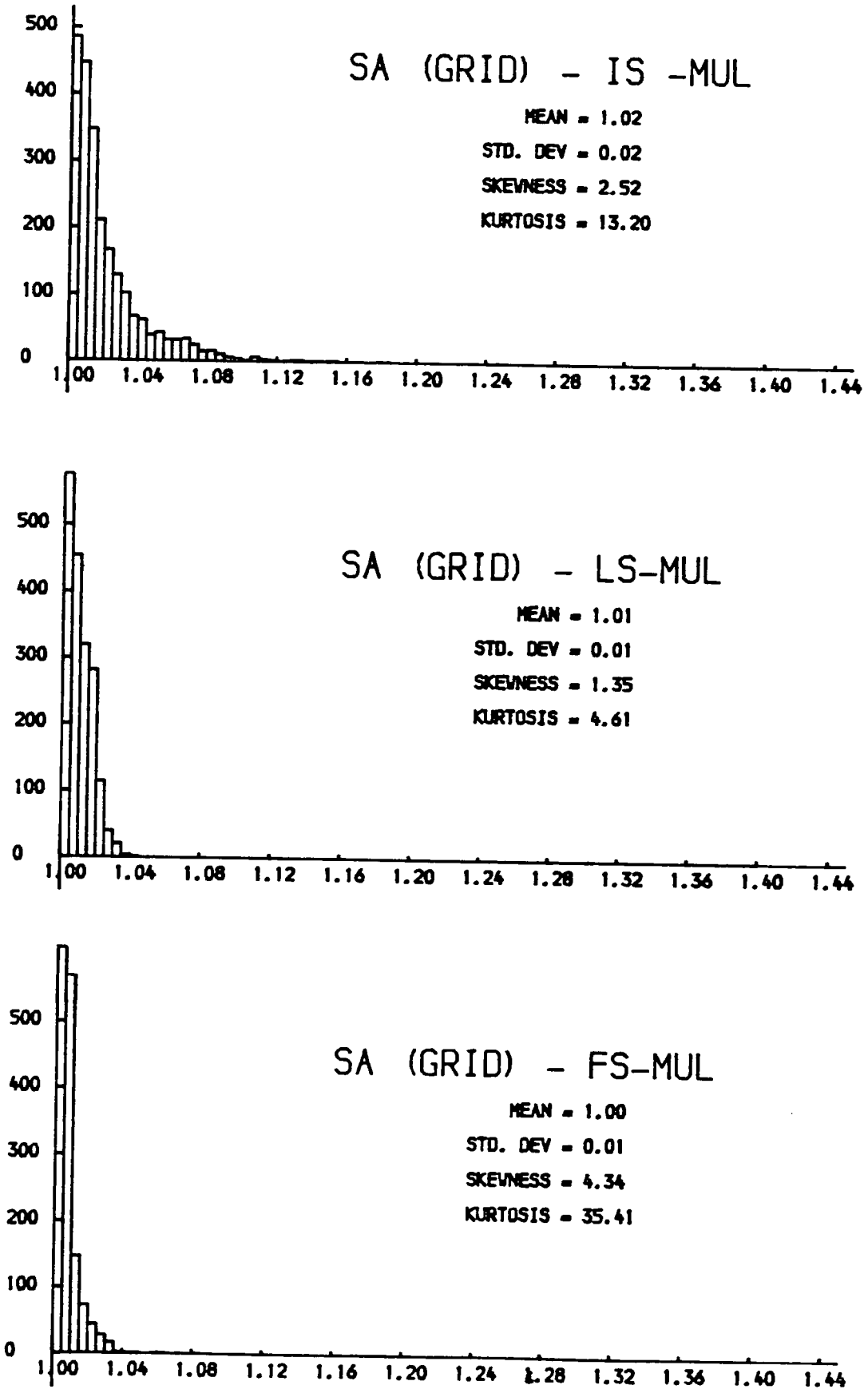


Figure 4.7 Surface area analysis distribution - range of surface types.

descriptor was being investigated. In such a situation it is important to consider the untransformed distributions since the statistics are in the true surface units. More importantly, untransformed skewness is an important aid to interpretation, emphasising some of the problems and limitations of the technique. Similarly, untransformed kurtosis may be helpful although its correlation with skewness must be considered. Untransformed skewness is also a useful indicator as to whether the means and standard deviations are possibly distorted and therefore unrepresentative, allowing more emphasis to be placed on medians and inter-quartile ranges.

b. The resultant statistics were to be used in further correlations with error analyses statistics. In such a situation, statistics derived from a more normalised, transformed distribution are desirable with skewness minimised and kurtosis normalised. This allows a full use of all statistics and should improve the strength of the interpretation and the significance of the subsequent correlations.

The effect of transformation must therefore be considered. In this respect, much consideration was given to the work of Evans (1979) in comparing transformations of skewed distributions of gradient since surface area is essentially a form of pseudo-gradient - the steeper the gradient, the greater the total surface area.

Various transformations were considered in an attempt to minimise the strong positive skewness of the surface area distributions. Following a brief literature search and empirical tests, several basic transformations were applied to a selection of the untransformed distributions. From the preliminary results, two transformations were specially developed (2 and 3) and two transformations recommended by Evans (4 and 5) were applied. The final transformations selected were.

1. no transformation;
2. $\log_{10}(1000.0 \cdot (sa - 0.99))$;
3. $\tan^{-1}(1000.0 \cdot (sa - 1.0))$;
4. $\log_{10}(10.0 \cdot (\tan(sa))$;
5. $\sqrt{(2.0 \cdot \sin(sa))}$;

where sa = the individual surface area values being transformed.

In order to compare the degree of deviation from normality as a result of transformation, Evans considered that,

> "skewness is without doubt the most important single statistic (followed by kurtosis) ... it is desirable to find a transformation which minimises skewness. Given low skewness, normal kurtosis is the next <u>desideratum</u>" (Evans, 1979, 46).

Table 4.4 summarises the skewness and kurtosis for each transformation for each data set.

Evans also suggests that a fuller picture of deviation from normality is provided by a plot on cumulative probability paper. Class limits are plotted on one axis against cumulated percent frequency. It should be noted that, although the two tails of such a plot are important, probability paper 'stretches' both tails and therefore the importance of the

DATA	1		2		3		4		5	
	SKEW	KURT	SKEW	KURT	SKEW	KURT	SKEW	KURT	SKEW	KURT
I49	3.22	18.96	0.55	2.86	-3.49	16.10	4.24	32.02	2.40	11.00
I25	2.60	12.47	0.55	2.90	-3.51	15.58	2.91	15.26	2.20	9.43
I13	2.44	10.85	0.62	3.12	-3.11	12.53	2.60	11.95	2.20	9.25
IB-MUL	3.01	15.40	0.89	3.48	-3.17	15.21	3.26	17.82	2.66	12.39
IC-MUL	2.88	15.94	0.70	3.00	-3.05	13.63	3.21	19.28	2.46	12.06
IG-MUL	2.47	10.36	0.78	3.14	-3.05	12.53	2.61	11.41	2.25	8.91
IR-MUL	2.85	17.96	0.54	2.81	-2.85	11.01	3.33	24.56	2.34	12.12
IS-MUL	2.51	13.15	0.62	2.80	-3.02	12.64	2.83	17.08	2.14	9.55
L49	2.56	11.71	0.98	3.40	-1.47	4.33	2.65	12.47	2.40	10.51
L25	1.68	6.88	0.70	2.48	-1.02	2.71	1.72	7.18	1.60	6.34
L13	1.08	3.10	0.63	2.12	-0.67	1.85	1.09	3.12	1.06	3.05
LB-MUL	1.80	8.13	0.73	2.56	-0.64	1.75	1.85	8.55	1.70	7.41
LC-MUL	2.43	15.76	0.87	3.04	-0.83	2.10	2.57	17.68	2.22	13.02
LG-MUL	1.34	4.21	0.78	2.43	-0.55	1.66	1.35	4.28	1.31	4.09
LR-MUL	0.89	2.69	0.53	1.89	-0.55	1.62	0.90	2.72	0.87	2.64
LS-MUL	1.31	4.54	0.65	2.27	-0.74	1.85	1.34	4.65	1.27	4.33
F65	5.92	55.04	1.96	7.22	-1.01	2.89	8.03	93.92	4.71	32.21
F33	4.67	32.24	2.09	7.72	-0.67	2.07	4.92	36.02	4.30	27.17
F17	3.22	14.54	1.96	6.85	-0.60	2.01	3.26	14.81	3.15	14.02
FB-MUL	5.12	42.54	2.33	9.67	-0.67	2.19	5.37	47.22	4.86	36.75
FC-MUL	7.26	78.01	2.74	13.61	-0.39	1.76	7.63	85.64	6.79	67.74
FG-MUL	5.62	53.34	2.32	10.52	-0.42	1.77	5.85	57.53	5.42	48.03
FS-MUL	4.28	35.07	1.91	7.69	-0.51	1.86	4.45	37.90	4.18	31.61

Table 4.4 Summary of skewness and kurtosis for each transformation of surface area distribution for each data set.

extreme points should not be exaggerated. This technique was only used for the full grids (I49, L49 and F65 in Figure 4.8a,b and c) as it is clearly impractical to consider its use for all the data sets.

Finally, Evans argued against purely statistical comparisons to determine the optimum transformation since, in his opinion,

> "significance testing can be a red herring", (Evans, 1979, 6).

On the basis of his evidence, no such simplistic selection of optimal transformation was made.

On the basis of the skewness and kurtosis alone, the two transformations recommended by Evans seem to be unsatisfactory. While the square root sine transformation (5) produces marginally superior results to the use of no transformation, all logtan transformations have poorer statistics. This is supported by the plots which produce uneven concave curves, similar to those for no transformation, suggesting positive skewness.

While the two specially prepared transformations (log and arctan) are generally superior to no transformation, neither transformation is consistently the best. The log transformation has a lower (and hence 'superior') mean skewness (1.11) compared with the arctan transformation (-1.56). The log transformation is 'superior' where untransformed skewness is lower (all INCH and most LIAN results). However, where skewness is higher, the less powerful log transformation is unsuccessful compared with the more

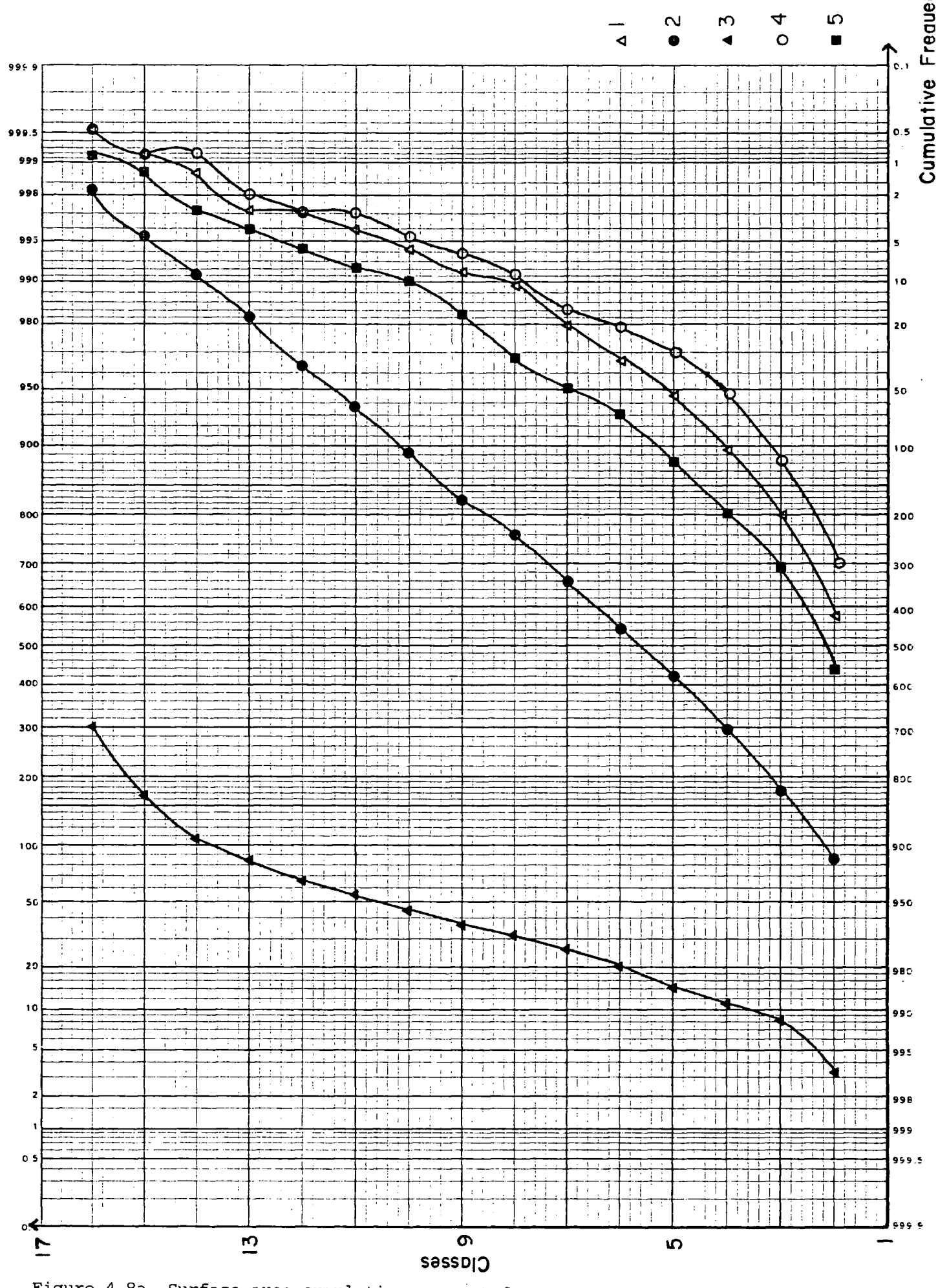


Figure 4.8a Surface area cumulative percent frequency distribution/class plot - INCH

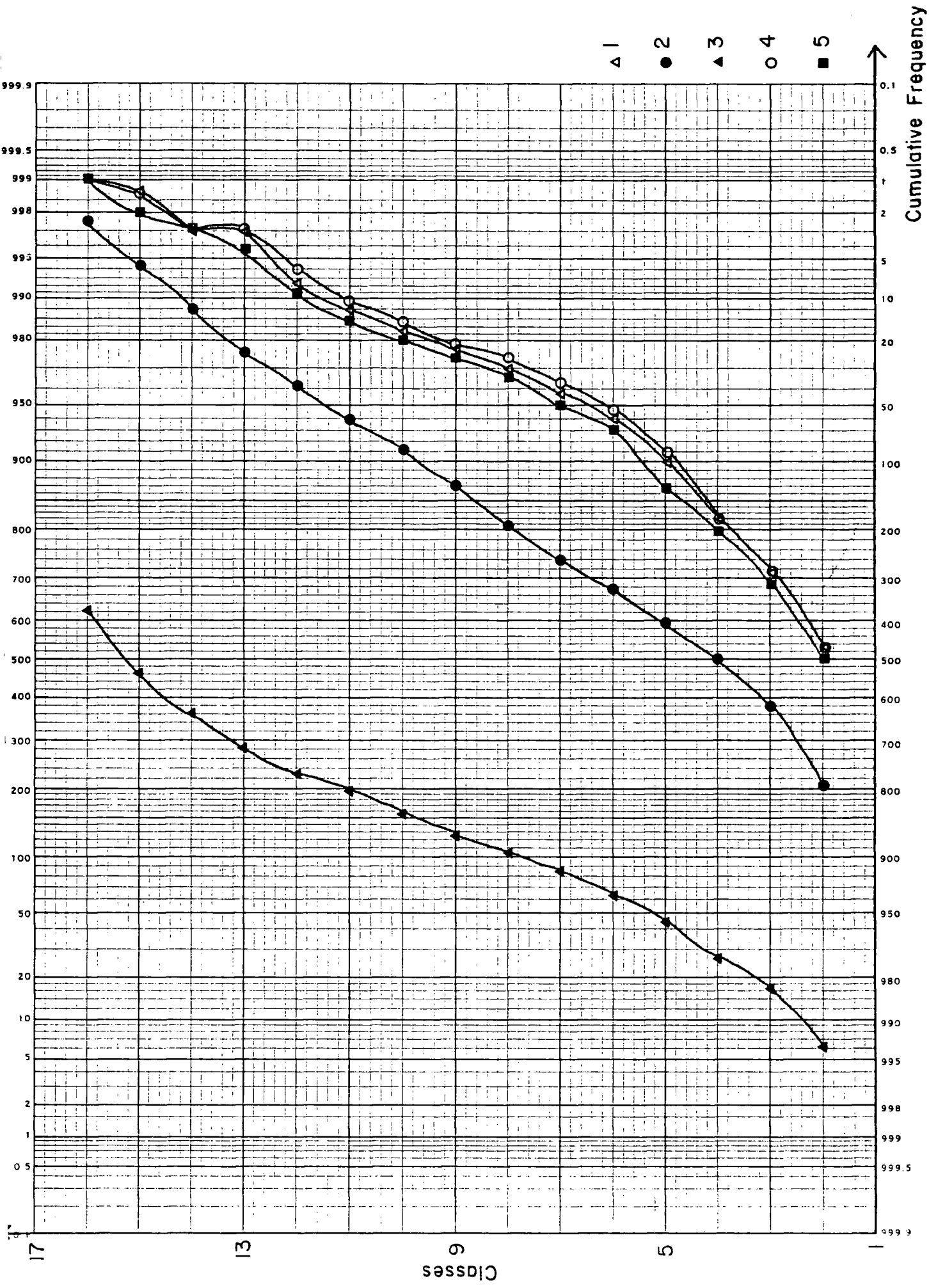


Figure 4.8b Surface area cumulative percent frequency distribution/class plot - LIAN

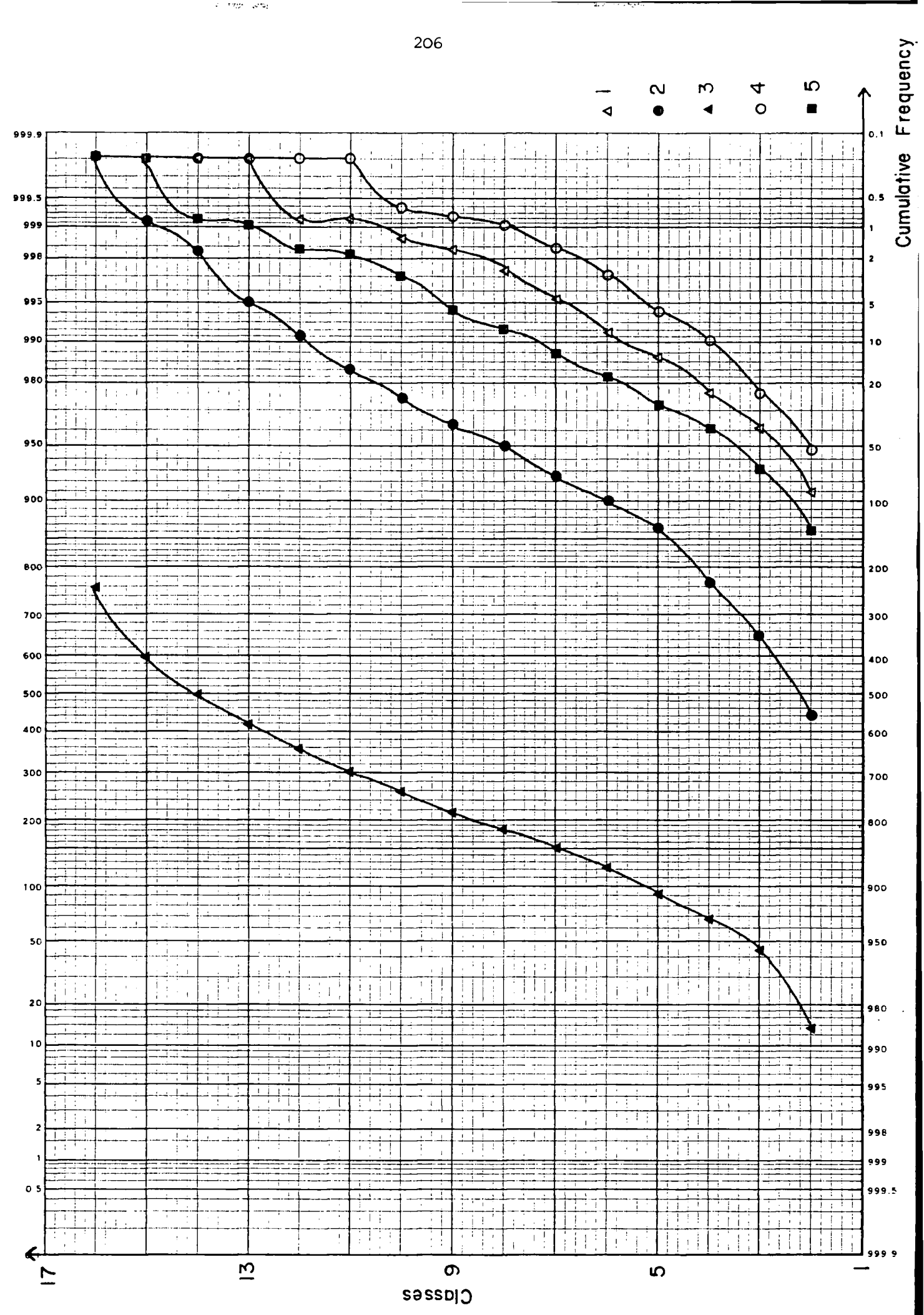


Figure 4.8c Surface area cumulative percent frequency distribution/class plot - FORV

powerful arctan transformation.

These conclusions are supported by the plots. The INCH and LIAN log transformation curves and the FORV arctan transformation curve are straighter and therefore more normalised. In each case, the alternative prepared transformation is S-shaped, which Evans suggests represents an unskewed kurtic distribution.

The results in general confirm that, while the normalisation of the untransformed distributions may be improved by transformation, there is no unique transformation which gives superior results. This casts doubts on the use of transformations in this particular situation, since some results may be improved at the expense of others in other data models. The lack of success of the 'recommended' transformations is probably as a result of the nature of the distribution and the fact that greater care was obviously taken over 'tuning' the other prepared transformations. Additionally, it is apparent that the changing patterns of median and mean, and inter-quartile and standard deviation, are similar enough to be considered identical in the following discussion.

It was decided to utilise the log transformation, since unlike the arctan transformation, it consistently improved the skewness and kurtosis. All statistics are summarised in Tables 4.5a,b and c for all data sets using no transformation (1) and using the log transformation (2). From the statistics, it is apparent that two important factors exist in the use of this descriptor namely, grid, resolution and surface type. The

DATA	MIN	MAX	MEAN	MEDIAN	I/Q	S.DEV	SKEW	KURT
I49	1.000	1.377	1.032	1.019	0.033	0.039	3.220	18.963
I25	1.000	1.217	1.026	1.017	0.026	0.029	2.597	12.473
I13	1.000	1.134	1.020	1.013	0.018	0.021	2.444	10.845
IB-MUL	1.000	1.183	1.019	1.012	0.018	0.023	3.009	15.386
IC-MUL	1.000	1.203	1.020	1.012	0.021	0.023	2.884	15.939
IG-MUL	1.000	1.172	1.020	1.012	0.018	0.023	2.468	10.362
IR-MUL	1.000	1.242	1.020	1.014	0.022	0.022	2.853	17.964
IS-MUL	1.000	1.248	1.021	1.013	0.022	0.023	2.507	13.152

1) Using no transformation.

DATA	MIN	MAX	MEAN	MEDIAN	I/Q	S.DEV	SKEW	KURT
I49	1.002	2.588	1.502	1.466	0.442	0.306	0.553	2.855
I25	1.006	2.355	1.464	1.429	0.358	0.274	0.552	2.900
I13	1.012	2.160	1.409	1.364	0.303	0.241	0.624	3.123
IB-MUL	1.001	2.286	1.379	1.335	0.341	0.248	0.886	3.482
IC-MUL	1.000	2.329	1.390	1.346	0.374	0.255	0.697	2.997
IG-MUL	1.000	2.260	1.386	1.344	0.332	0.256	0.776	3.140
IR-MUL	1.000	2.402	1.394	1.375	0.385	0.254	0.538	2.810
IS-MUL	1.001	2.411	1.405	1.359	0.370	0.260	0.623	2.796

2) Using a logarithmic transformation.

Table 4.5a Statistical summary of the effects of a logarithmic transformation within surface area analysis (all INCH grids).

DATA	MIN	MAX	MEAN	MEDIAN	I/Q	S.DEV	SKEW	KURT
L49	1.000	1.102	1.011	1.006	0.013	0.013	2.558	11.709
L25	1.000	1.061	1.008	1.004	0.012	0.009	1.679	6.877
L13	1.000	1.025	1.006	1.004	0.008	0.007	1.077	3.097
LB-MUL	1.000	1.071	1.007	1.004	0.012	0.008	1.799	8.131
LC-MUL	1.000	1.109	1.007	1.004	0.010	0.009	2.434	15.755
LG-MUL	1.000	1.038	1.006	1.003	0.009	0.007	1.339	4.210
LR-MUL	1.000	1.027	1.005	1.003	0.010	0.006	0.892	2.694
LS-MUL	1.000	1.054	1.007	1.004	0.011	0.008	1.313	4.539

1) Using no transformation.

DATA	MIN	MAX	MEAN	MEDIAN	I/Q	S.DEV	SKEW	KURT
L49	1.001	2.051	1.256	1.195	0.324	2.214	0.981	3.404
L25	1.000	1.851	1.218	1.158	0.314	0.190	0.704	2.481
L13	1.002	1.543	1.177	1.137	0.259	0.163	0.631	2.120
LB-MUL	1.000	1.910	1.190	1.135	0.315	0.183	0.732	2.559
LC-MUL	1.000	2.074	1.189	1.141	0.290	0.178	0.874	3.040
LG-MUL	1.000	1.679	1.167	1.116	0.262	0.163	0.777	2.429
LR-MUL	1.000	1.567	1.158	1.113	0.279	0.148	0.531	1.888
LS-MUL	1.000	1.804	1.187	1.145	0.309	0.174	0.653	2.274

2) Using a logarithmic transformation.

Table 4.5b Statistical summary of the effects of a logarithmic transformation within surface area analysis (all LIAN grids).

DATA	MIN	MAX	MEAN	MEDIAN	I/Q	S.DEV	SKEW	KURT
F65	1.000	1.426	1.011	1.004	0.009	0.025	5.918	55.035
F33	1.000	1.153	1.007	1.002	0.006	0.014	4.670	32.241
F17	1.000	1.047	1.005	1.002	0.004	0.008	3.220	14.535
FB-MUL	1.000	1.127	1.004	1.002	0.004	0.008	5.118	42.537
FC-MUL	1.000	1.128	1.004	1.002	0.004	0.008	7.255	78.013
FG-MUL	1.000	1.094	1.003	1.002	0.004	0.006	5.616	53.336
FS-MUL	1.000	1.084	1.004	1.002	0.004	0.005	4.275	35.065

1) Using no transformation.

DATA	MIN	MAX	MEAN	MEDIAN	I/Q	S.DEV	SKEW	KURT
F65	1.000	2.639	1.221	1.131	0.254	0.247	1.963	7.217
F33	1.000	2.212	1.166	1.096	0.189	0.199	2.091	7.717
F17	1.001	1.753	1.131	1.082	0.137	0.153	1.956	6.852
FB-MUL	1.000	2.137	1.127	1.080	0.135	0.148	2.334	9.667
FC-MUL	1.000	2.138	1.109	1.062	0.131	0.137	2.739	13.610
FG-MUL	1.000	2.016	1.104	1.064	0.118	0.122	2.324	10.518
FS-MUL	1.000	1.972	1.114	1.074	0.141	0.123	1.908	7.689

2) Using a logarithmic transformation.

Table 4.5c Statistical summary of the effects of a logarithmic transformation within surface area analysis (all FORV grids).

former, grid resolution, should be examined first as it has a bearing on general surface comparisons, since subsequent comparisons involve comparing grids of different size and resolution. It is evident from Figure 4.6 and Table 4.5 that for each surface, as grid resolution decreases and thus the number of grid cells decreases, so all the quoted statistics decrease (with the exception of skewness and kurtosis in the log transformation statistics). From the histograms, it is also apparent that the percentage decline is proportional across the histogram, but that the lower values numerically decrease most rapidly. Thus the surface is undergoing a smoothing or flattening effect, exemplified by the 1.005-1.010 class of Figure 4.6. This is entirely understandable, since by decreasing the grid resolution, the total gradient may not be greater than the greatest component gradient, see Figure 4.9. Therefore, the maximum, mean, median, inter-quartile range and standard deviation must all decrease. Notably, skewness and kurtosis do not decrease so rapidly since the shape of the histogram is still preserved. This is borne out in the transformed statistics where these values increase in INCH and FORV data. Overall, it is important that all surfaces respond similarly, although INCH has a slower numerical decline due to a few extreme outliers caused by the scarp-slope in the North-West of the area.

Comparing surfaces, once again distinct patterns emerge from the untransformed statistics and the histograms. FORV data gives consistently smaller means and thus, since there is a minimum surface area for a cell (1.0), much greater values of skewness and kurtosis. LIAN data, while giving similar

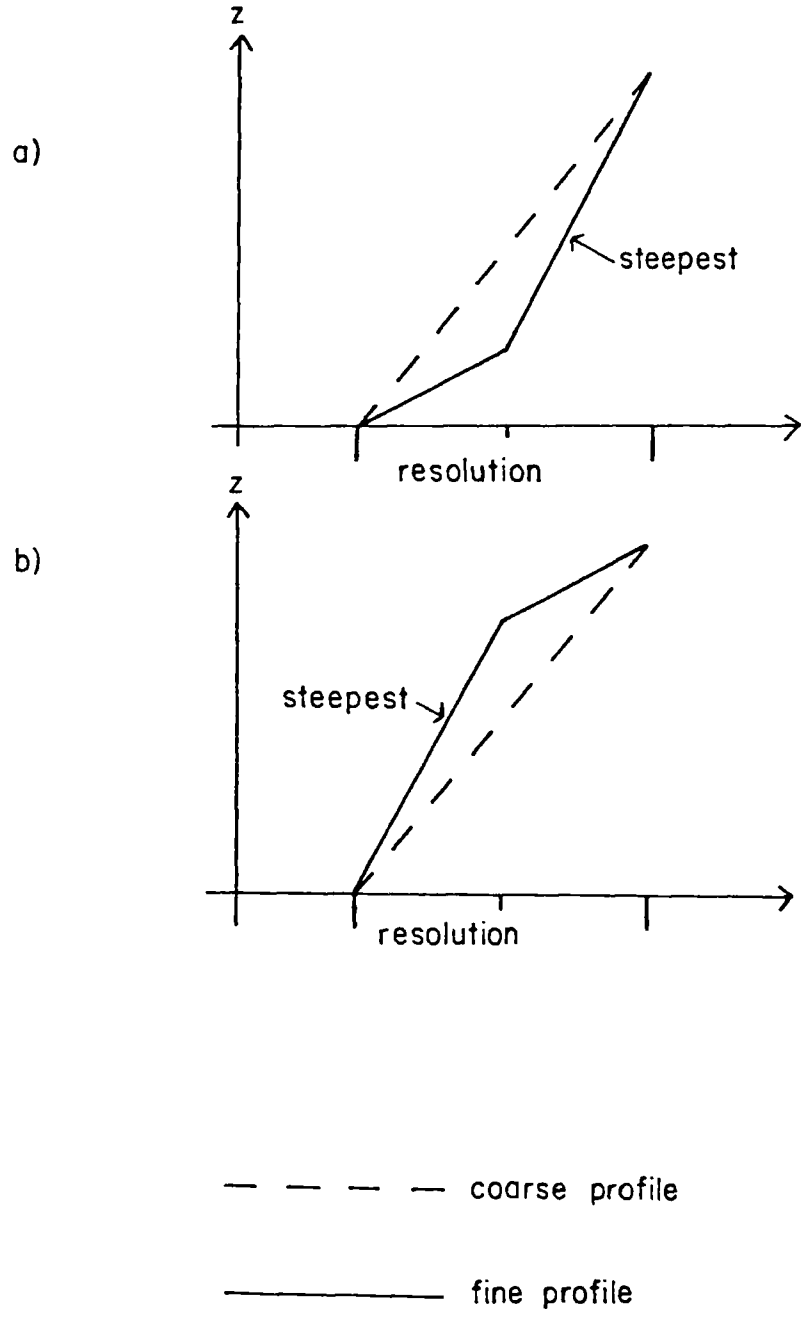


Figure 4.9 Proof that the gradient of a coarse resolution grid cell cannot be greater than the maximum gradient of its component cells.

standard deviations to FORV data, produce slightly higher means and, with the exception of CONTOUR data, smaller values for skewness and kurtosis. LIAN is most easily interpreted since the surface is a smooth continuous slope offering a more even distribution of slopes and thus of cell surface-areas. The larger 1.000 - 1.005 class is undoubtedly caused by the large flat river basin area within the model and, without this, perhaps the skewness would have been nearer zero. Kurtosis, with values for L13 (3.103) and LR-MUL (2.760) both near the important 3.0 index (of normal kurtic distribution), implies a near perfect normal distribution.

INCH results suggest a rougher and more varied terrain than either of LIAN or FORV. Mean values of approximately 1.020 (INCH) as opposed to 1.004 (FORV) and 1.006 (LIAN) are a result of both a wider range of slopes and thus cell surface areas, but also the existence of some extreme values, for example between 1.16 and 1.20, and 1.24 on the IS-MUL histogram (Figure 4.7). These aspects also produce much higher standard deviations, while skewness lies between the almost normal distribution of LIAN and the markedly skewed distribution of FORV, and kurtosis is not as pronounced as that of FORV.

Thus, in summary, FORV data are distinguished by very high skewness and kurtosis and very low mean, suggesting very flat and smooth terrain. LIAN is characterised by a low mean and standard deviation but also by near normal distribution. INCH data have a much higher mean and standard deviation than have LIAN or FORV data. Thus it would appear that this descriptor

is only capable of detecting macro-topographic features and ignores micro-nuances of the different areas.

Additionally, while the different resolutions and number of cells were shown to have an effect on the descriptor, it is apparent that this would have little added implication for surface comparisons. If the results were standardised on the basis of the number of cells, INCH and LIAN would remain virtually unchanged, while the results for FORV would fall slightly, producing a more distinctively smaller mean and a skewness and kurtosis still larger but closer to INCH, but markedly greater than LIAN. Alternatively, if the standardisation was based on resolution, assuming INCH remains unchanged, FORV would have a considerably smaller mean but have skewness and kurtosis similar to INCH. LIAN would have a mean greater than FORV, but with a skewness and kurtosis virtually normal.

Concentrating on the ability of the surface area descriptor to distinguish the different data subsets, it is apparent from the previous discussion, Table 4.5 and Table 4.6 that - like spatial autocorrelation - no clear general pattern emerges within the statistics. The within-data surface consistency of untransformed maximum, mean median, inter-quartile and standard deviation statistics has been discussed. While the range of these individual statistics increases with transformation, no apparent consistency emerges across the surfaces (see Table 4.6 summarising the means and standard deviations). The greatest consistency occurs with untransformed skewness. In general, BREAKLINE and CONTOUR data

	TRANSFORMED MEAN			TRANSFORMED STANDARD DEVIATION			UN-TRANSFORMED SKEWNESS		
	INCH	LIAN	FORV	INCH	LIAN	FORV	INCH	LIAN	FORV
SMALLEST	BREAKLINE	RIVER	GRID	BREAKLINE	RIVER	GRID	GRID	RIVER	SCATTER
	GRID	GRID	CONTOUR	RIVER	GRID	SCATTER	SCATTER	SCATTER	GRID
	CONTOUR		SCATTER	CONTOUR	SCATTER	CONTOUR	RIVER	GRID	BREAKLINE
	RIVER	SCATTER	BREAKLINE	GRID	CONTOUR	BREAKLINE	CONTOUR	BREAKLINE	CONTOUR
		CONTOUR		SCATTER	BREAKLINE		BREAKLINE	CONTOUR	
LARGEST	SCATTER	BREAKLINE							

Table 4.6 Within-model comparison of some surface area statistics for regular grids.

produce more extreme skewness and kurtosis values. RIVER data produce either low or very high values and GRID and SCATTER data produce low-to-average results. The latter do have a more even data distribution, therefore extreme surface slopes have been removed as large interpolated areas (the reader is reminded that these gridded data sets have been interpolated using multiquadric analysis) will have similar homogeneous surface characteristics. The BREAKLINE and CONTOUR distributions are not so even, therefore small and large patches of homogeneous surface characteristics will exist. If some of these patches have extreme gradients, for example in the large blow-out area (SW of centre) in FORV and the scarp slope (centre N) in INCH, then a greater range of values may exist, although due to the larger smooth patches, the distribution may be more skewed. The RIVER data sets are more difficult to explain, although the discrepancy between INCH and LIAN is possibly related to the original model surface. LIAN-RIVER data have relatively a much smaller range of z values than INCH and therefore the interpolated surface is smoother, resulting in much smaller means, standard deviations, skewness and other statistics.

More interestingly, the multiquadric interpolation appears to have over-smoothed and flattened the data. In general, means and standard deviations are considerably smaller than the comparative I49, L49 and F65 reference grids. This has had an additional effect of making the distribution more normal and thus generally smaller skewness and kurtosis values are found than in the 'correct' or 'ground truth' data.

4.4.3.2 Irregular Data

As a result of the negligible increase in information gained from the transformation investigation within gridded surface area analysis, this aspect was not considered within irregular data surface area analysis. While in retrospect, with such high skewnesses this may have been useful, it was not possible to perform this investigation due to the practical computing limitations of accessing the package. The untransformed results of using irregular data in 'DTSA' show that the descriptor produces much more varied results with point data. This is easily illustrated by considering standard deviation (see Table 4.7 and Figure 4.10). In general, there is a greater spread of values across the histogram and thus the standard deviation for point data, even when standardised for the number of cells, is considerably greater than any of the relevant 'SA' grid standard deviations. This wider distribution is a direct result of extreme values caused by small, irregular, steep-sided cells which delineate local nuances of the terrain. These small cells have as much significance as every other cell (including very large, perhaps over-smooth cells), resulting in distortion of the histogram and parameters.

While most of the parameters are generally larger than the equivalent 'SA' parameters, nevertheless similar patterns emerge. Larger data sets, for example IB3 and IC3, generate larger values for all the parameters than do the smaller data sets (IB and IC). The FORV point data set gives consistently larger means and standard deviations although, unlike gridded

DATA	MEAN	STD.DEV	SKEW	KURT	CELLS	RESO
I49 (GRID)	1.032	0.039	3.220	18.963	2304	30
I25 (GRID)	1.026	0.029	2.597	12.473	576	60
I13 (GRID)	1.020	0.021	2.444	10.845	169	120
IB-MUL (GRID)	1.019	0.023	3.009	15.386	2304	30
IC-MUL (GRID)	1.020	0.023	2.884	15.939	2304	30
IG-MUL (GRID)	1.020	0.023	2.468	10.362	2304	30
IR-MUL (GRID)	1.020	0.022	2.853	17.964	2304	30
IS-MUL (GRID)	1.021	0.023	2.507	13.152	2304	30
IB	1.052	0.143	8.921	102.043	603	
IC	1.042	0.082	9.409	128.618	535	
IR	1.061	0.185	8.173	87.888	535	
IS	1.060	0.168	8.557	97.526	537	
IB3	1.053	0.113	7.777	94.877	1097	
IC3	1.046	0.136	15.653	321.602	1096	
L49 (GRID)	1.011	0.013	2.558	11.709	2784	15
L25 (GRID)	1.008	0.009	1.679	6.877	696	30
L13 (GRID)	1.006	0.007	1.077	3.097	182	60
LB-MUL (GRID)	1.007	0.008	1.799	8.131	2784	15
LC-MUL (GRID)	1.007	0.009	2.434	15.755	2784	15
LG-MUL (GRID)	1.006	0.007	1.339	4.210	2784	15
LR-MUL (GRID)	1.005	0.006	0.892	2.694	2784	15
LS-MUL (GRID)	1.007	0.008	1.313	4.539	2784	15
LB	1.024	0.049	9.882	154.788	558	
LC	1.041	0.194	14.183	228.758	585	
LR	1.019	0.058	15.939	315.859	542	
LS	1.007	0.015	5.723	52.507	585	
LB3	1.035	0.091	14.088	282.608	1184	
LC3	1.049	0.310	13.954	224.741	1026	
F65 (GRID)	1.011	0.025	5.918	55.035	4096	10
F33 (GRID)	1.007	0.014	4.670	32.241	1024	20
F17 (GRID)	1.005	0.008	3.220	14.535	256	40
FB-MUL (GRID)	1.004	0.008	5.118	42.537	4096	10
FC-MUL (GRID)	1.004	0.008	7.255	78.013	4096	10
FG-MUL (GRID)	1.003	0.006	5.616	53.336	4096	10
FS-MUL (GRID)	1.004	0.005	4.275	35.065	4096	10
FB	1.014	0.032	6.165	55.848	554	
FC	1.014	0.037	6.108	48.819	554	
FS	1.011	0.021	3.695	19.621	554	
FB3	1.025	0.057	10.640	190.004	1170	
FC3	1.018	0.051	7.903	88.914	1048	

Table 4.7 Comparative results of point and gridded surface area analyses using untransformed statistics.

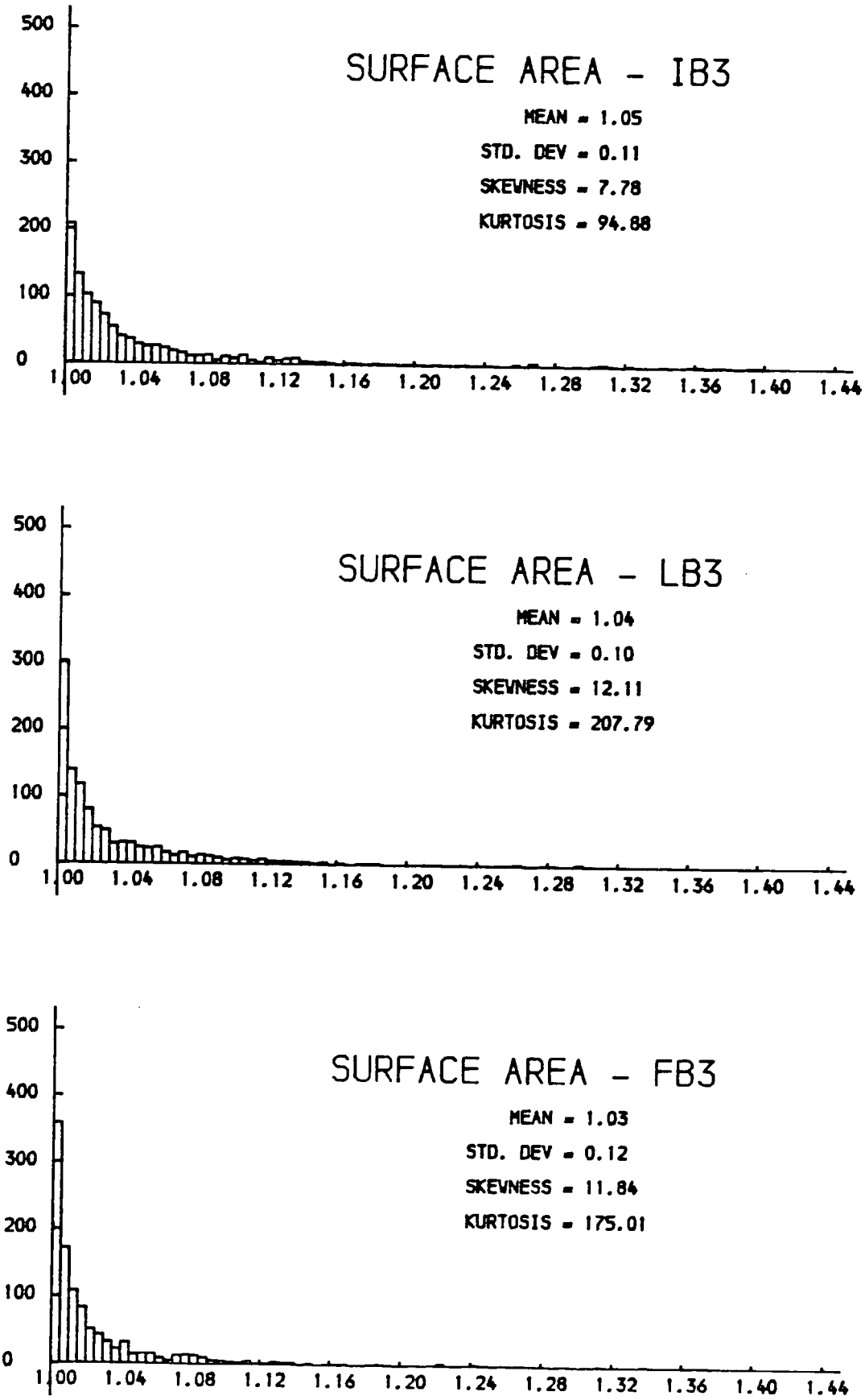


Figure 4.10 Surface area analysis distribution - point data

parameters, the skewness and kurtosis values are considerably smaller than the similar INCH data set - therefore inconsistency. The LIAN point data set gives perhaps the most varied results and the parameters provide a very tenuous consistency. Means and standard deviations are distinctively larger than with the FORV data and smaller than INCH data, although skewness and kurtosis seem totally unpredictable and the results difficult to comprehend. Skewness and kurtosis range widely the latter from 52 to 315, thus the results seem highly dependent on data. Finally INCH, with the exclusion of the increased CONTOUR subset (IC3), follows a similar pattern to FORV data only with higher values. Thus, in summary, the three surfaces are still distinguishable and seem to be influenced marginally by increases in data set sizes.

The differentiation of data subsets is much more difficult. BREAKLINE and CONTOUR data seem to produce similar results within any surface type, with the exclusion of IC3 and FB3. RIVER data, while producing varied means and standard deviations, produce consistently smaller values for skewness and kurtosis than any other data type. SCATTER data produce more varied results, and thus no definitive pattern emerges.

Comparing the DTSA and SA results and thus the ability of point data to match gridded data, it is apparent that both methods do not produce identical results. While both methods create an ability to define each surface type from the parameters, this is often for differing reasons. Mean and standard deviations produced by the two programs broadly follow similar patterns, but vary in magnitude. Skewness and

kurtosis are completely different and kurtosis seems especially highly variable with point data - as a result of extreme outliers and poor data distributions. Within any one surface, similar patterns emerge for data types. Magnitudes may be radically increased, but similar patterns are at work for comparing skewness and kurtosis for interpolated grids against point data. Thus it is suggested that point and gridded data surface area parameters do react similarly, although the use of interpolated grids appears to filter out some of the extremeness from the point data parameters. However, this extremeness may well be a meaningful advantage of point data.

While the four parameters form a fairly comprehensive descriptor of the terrain, their global consistency is limited. Ideally, excluding any scale difference between the grid and point results, similar patterns should emerge for the data sets. While these do occur when considering certain items, they are by no means universal to the method, in terms of grid data producing exactly the same response as point data, and thus the consistency must be considered poor.

4.4.4 Dispersion Vectors

The use of dispersion vectors for measuring surface roughness was suggested by Hobson (1967) and discussed by Turner and Miles (1968). Calculation of vector dispersion within this thesis involved the implementation of a program 'VECTOR', a modification of the similarly-named Hobson program. The program requires the input of a regular grid of

heights, from which is derived a set of triangular planar surfaces defined by adjacent grid nodes. The parameters, total vector strength (R1); standardised vector strength (R); dispersion factor (K) and terrain variability measure (V) are subsequently derived from these triangular surfaces.

On considering the R1, R, K and V parameters produced from the INCH, FORV and LIAN observed and interpolated grids (see Table 4.8), two main factors were found to be contributory to the results, namely, grid size/resolution and surface type. Over the three data surfaces, decreasing the grid resolution and thus decreasing the number of grid nodes has the effect of increasing both the standardised vector strength (R) and dispersion factor (K). In addition, this decreases the variability factor (V) which is computed from the inverse of log (K), although the effect is less marked. It is therefore essential that some knowledge of grid resolution is required before different samples are compared by this method.

In general - even after making allowances for grid standardisation - the variability factors produced for each surface are consistent within surface type, and markedly different between surface types. FORV data produce low V values (approximately 12) suggesting a relatively smooth surface ranging somewhere between Turner's 'karst plain' and 'plateau' terrain types (Table 4.1). LIAN data suggest a rougher surface (approximately 41) equivalent to Turner's 'hill' classification. Finally INCH data produce the roughest V value (approximately 105) which falls outside the 'escarpment' classification. Since the standard deviations of

DATASET	SIZE	RANGE	R1	R	K	V
SMOOTH ROUGH				HIGH LOW	HIGH LOW	
I49	49x49	180.9	4470	0.9705	33.90	118.2
I25	25x25	180.9	1140	0.9753	40.40	112.6
I13	13x13	178.3	286	0.9805	51.23	104.3
IB-MUL	49x49	179.9	4523	0.9819	55.13	103.3
IC-MUL	49x49	178.6	4519	0.9811	52.86	103.6
IG-MUL	49x49	179.4	4520	0.9812	53.12	103.9
IR-MUL	49x49	178.3	4519	0.9809	52.37	103.7
IS-MUL	49x49	180.0	4514	0.9800	49.88	106.0
L49	49x59	91.7	5316	0.9896	96.22	46.2
L25	25x30	91.5	1333	0.9919	123.42	43.7
L13	13x15	86.6	334	0.9939	163.98	39.1
LB-MUL	49x59	90.4	5525	0.9931	144.86	41.8
LC-MUL	49x59	92.3	5526	0.9931	145.89	42.6
LG-MUL	49x59	89.1	5532	0.9943	174.52	39.7
LR-MUL	49x59	90.3	5535	0.9948	191.63	39.5
LS-MUL	49x59	90.1	5527	0.9934	150.89	41.3
F65	65x65	30.1	7846	0.9896	96.38	15.1
F33	33x33	29.6	2034	0.9933	148.63	13.6
F17	17x17	29.6	510	0.9955	219.49	12.6
FB-MUL	65x65	29.1	8145	0.9957	232.21	12.3
FC-MUL	65x65	27.6	8151	0.9964	277.37	11.3
FG-MUL	65x65	29.0	8154	0.9968	313.20	11.6
FS-MUL	65x65	28.3	8152	0.9965	284.32	11.5

Table 4.8 Results of vector dispersion analysis

these values within any surface type are remarkably low, this appears to give an excellent single surface roughness descriptor.

The variability factor is not the only parameter computed and, while it clearly differentiates surface type, its descriptive power is increased by the use of R and K values. INCH interpolated data produce distinctive R and K values (means of 0.9810 and 52.7). LIAN (0.9937 and 161.6) and FORV (0.9963 and 276.8) produce more similar R statistics, although after standardisation, both R and K values increase for FORV and thus the two surfaces become more distinct, the vector strengths (R) and dispersion factors (K) differing more significantly.

Considering the data subsets, no significant patterns emerge within the surfaces (see Tables 4.8 and 4.9). For example, CONTOUR, GRID and RIVER data produce very similar results in the INCH surface although CONTOUR results are similar to BREAKLINE and SCATTER results in LIAN, and differ from the other data sets in FORV. Similarly, SCATTER data produce distinctive results with the INCH surface, is similar to BREAKLINE and CONTOUR in LIAN and similar to GRID results in FORV. Thus, as a possible result of either too much interpolation 'noise' or a basic lack of consistent significant differences, this descriptor cannot distinguish between the different data subsets satisfactorily.

4.4.5 Pointwise Polynomials

An integral part of the "Integrated system of terrain

	K			V		
	INCH	LIAN	FORV	INCH	LIAN	FORV
LARGE	BREAKLINE	RIVER	GRID	BREAKLINE	RIVER	CONTOUR
				CONTOUR	GRID	
	GRID	GRID	SCATTER	RIVER		SCATTER
	COUNTOUR		CONTOUR	GRID	SCATTER	GRID
	RIVER	SCATTER			BREAKLINE	
		CONTOUR				
SMALL	SCATTER	BREAKLINE	BREAKLINE	SCATTER	CONTOUR	BREAKLINE

Table 4.9 Within-model comparisons of K and V VECTOR statistics

analysis and slope mapping" discussed by Evans (1979) is the program 'GEOD', described and documented by Young (1978). This program was accessed at NUMAC (see A1.3.3) and used to produce the pointwise polynomial decriptors (see Tables 4.10a-e) as discussed earlier. Scattergrams, histograms and hachure maps are also produced by this program to illustrate the tabular information but, since they duplicate much of the relevant statistics, they will not be considered here, especially since they form a very bulky form of statistical presentation. Instead, the limited discussion will revolve around the summary tables. It should be noted that the program, in computing kurtosis values, subtracts a constant value of 3.0 from the computed value, so that the Gaussian expectancy is zero. All units are as specified in the relevant tables. The summary statistics will be reviewed in order to elucidate their geomorphic meaning.

This form of surface characteristic statistic may have benefitted from applying transformations to the parameter distributions, however these were not applied. As noted earlier, Evans (1979) suggested that there is no unique transformation which may applied to give consistently improved results; such an investigation would have been very lengthy and would not have been central to this thesis. Finally, as suggested by results in the section on surface analysis, the resultant benefits seem likely to be small and would remove the ability to consider untransformed skewness as a valuable statistical descriptor.

Altitude parameters (Table 4.10a) show few important

AREA	MEAN	S.D.	SKEWNESS	KURTOSIS	MAX	MIN
I49	160.7	43.6	-0.275	-0.757	252	71
I25	159.9	42.8	-0.266	-0.740	250	73
I13	167.8	19.2	-0.837	-0.523	213	143
IB -MUL	159.6	41.9	-0.298	-0.686	246	69
IC -MUL	158.8	42.0	-0.291	-0.749	246	70
IG-MUL	159.1	41.8	-0.268	-0.711	250	73
IR-MUL	150.4	43.5	-0.109	-0.948	246	71
IS -MUL	160.0	41.7	-0.285	-0.687	250	71
L49	84.1	20.8	1.116	0.181	151	63
L25	84.3	21.1	1.121	0.189	151	64
L13	83.0	19.3	1.097	0.125	140	65
LB -MUL	83.2	20.9	1.116	0.205	151	63
LC-MUL	83.9	20.4	1.116	0.198	146	59
LG-MUL	83.9	20.7	1.120	0.175	150	64
LR-MUL	83.2	20.7	1.197	0.520	152	64
LS-MUL	83.8	20.9	1.113	0.200	152	64
F65	38.8	5.1	0.740	0.810	55	26
F33	38.9	5.1	0.730	0.755	55	26
F17	38.9	4.9	0.740	0.775	54	28
FB -MUL	39.7	5.3	0.380	-0.194	55	26
FC-MUL	39.1	4.8	0.690	0.780	54	27
FG -MUL	38.9	5.0	0.690	0.825	55	26
FS-MUL	39.3	5.0	0.410	0.544	54	26

Table 4.10a Results of pointwise polynomials
ALTITUDE (metres)

features of the landscape. Evans suggests that mean altitude is of use purely as a location descriptor. Alternatively, variability is represented by standard deviation, and this shows two important features. Variability varies with scale, thus the coarser resolution DEMs have a lower standard deviation, although this is marginally so with FORV. Additionally, with the minor exceptions of FORV-BREAKLINE and INCH-RIVER, the other surface patterns are remarkably consistent, suggesting INCH data sets (42) are more variable than LIAN (20) and FORV (5).

The skewness and kurtosis values suggest more about the distribution within the altitude matrix. INCH has weakly negative skewness and kurtosis values, suggesting a slight dominance of higher values and thus a plateau effect. LIAN has small positive skewness caused by a predominance of low values in the river basin, and an almost normal kurtosis (with the notable exception of river data). FORV data show a smaller skewness, suggesting a slight inselberg appearance stressed by the strongest kurtosis values. Little data set discrimination may be established using altitude.

Gradient statistics (Table 4.10b) show a similar two-feature appearance, with a general absolute decrease for all parameters as a direct result of decreasing resolution. In addition, the three surfaces may be easily discerned. Evans (1979, 44) proposed that "mean gradients increase in line with variability of altitude". In accordance with this, INCH shows most variability, while LIAN is marginally more variable than FORV. This is emphasised on considering standard deviation,

AREA	MEAN	S.D.	SKEWNESS	KURTOSIS	MAX	MIN
I49	10.2	5.9	0.964	1.157	38	0
I25	9.1	5.0	0.898	0.708	28	0
I13	3.7	2.0	0.697	-0.151	9	0
IB -MUL	9.3	4.9	0.991	1.145	31	0
IC -MUL	9.5	5.1	0.892	0.919	32	0
IG-MUL	9.4	5.1	0.904	0.803	28	0
IR-MUL	9.7	5.1	0.611	0.299	33	0
IS -MUL	9.7	5.2	0.697	0.192	32	0
L49	5.5	3.9	0.546	-0.685	20	0
L25	5.0	3.6	0.351	-1.158	14	0
L13	4.6	3.5	0.254	-1.275	11	0
LB -MUL	5.2	3.7	0.351	-1.042	19	0
LC -MUL	5.2	3.6	0.422	-0.870	16	0
LG-MUL	4.9	3.6	0.456	-0.991	14	0
LR-MUL	4.7	3.4	0.246	-1.315	12	0
LS-MUL	5.2	3.7	0.300	-1.115	15	0
F65	4.5	4.1	2.182	6.697	31	0
F33	3.7	2.9	1.577	3.776	18	0
F17	3.0	1.8	0.714	0.025	8	0
FB -MUL	4.1	3.0	1.822	5.150	24	0
FC-MUL	3.7	2.9	1.872	5.890	23	0
FG -MUL	3.6	2.6	1.452	3.517	20	0
FS-MUL	3.9	2.6	1.077	1.460	18	0

Table 4.10b Results of pointwise polynomials
GRADIENT (degrees)

although the magnitudes are more marginal. Skewness of gradient is positive in all cases suggesting a dominance of lower gradients. FORV has the highest skewness, suggesting a very flat area with a few sudden rises suggested by the higher kurtosis values (although one should be wary of reading too much into the kurtosis statistic due to the higher skewness). LIAN, with low positive skewness and negative kurtosis, reflects the more even distribution of slope across the surface, with the skewness a result of the presence and location of the river basin, and kurtosis a result of the bimodal gradient distribution caused by the two different terrain linear slopes (see Figure 3.3c). Finally, INCH has a moderate positive skewness and kurtosis resulting from the plateau effect with a preponderance of tops and valley floors over the valley sides. In summary, little consistency exists between data sets, although all of the parameters may be used to differentiate surface type.

Aspect (Table 4.10c) produces perhaps the most distinct consistency. Mean values, which naturally vary from 0° - 360°, are of least importance since the orientation of the model is arbitrary. Nevertheless, INCH (75) may be considered predominantly ENE facing LIAN (208) SSW facing, and FORV (170) SSE. More important are vector strengths - INCH has the most varied orientation (0.29), FORV is marginally less varied (0.33), but LIAN is very directionally orientated (0.50) - essentially only a SSW facing hill-slope. This surface differentiation becomes remarkably consistent when 'gradient weighted' values are considered, with strengths identical for any data set within a particular surface.

AREA	UNIT VECTORS		GRADIENT WEIGHTED		NO.	RESO
	MEAN	STRENGTH	MEAN	STRENGTH		
I49	67.9	0.25	61.5	0.05	2209	30
I25	70.9	0.30	63.7	0.05	576	60
I13	99.5	0.31	121.5	0.03	121	120
IB -MUL	75.6	0.29	65.3	0.05	2209	30
IC -MUL	78.6	0.29	64.7	0.05	2209	30
IG-MUL	75.6	0.28	65.2	0.05	2209	30
IR-MUL	77.4	0.29	64.8	0.05	2209	30
IS -MUL	75.7	0.29	65.0	0.05	2209	30
L49	205.4	0.49	209.0	0.07	2679	15
L25	206.2	0.51	209.2	0.07	696	30
L13	198.6	0.53	207.3	0.07	143	60
LB -MUL	204.3	0.50	208.5	0.07	2679	15
LC -MUL	197.4	0.49	208.3	0.07	2679	15
LG-MUL	203.3	0.50	209.3	0.07	2679	15
LR-MUL	202.3	0.57	209.7	0.07	2679	15
LS-MUL	204.1	0.50	209.2	0.07	2679	15
F65	168.5	0.33	174.0	0.02	3969	10
F33	168.5	0.36	171.5	0.02	961	20
F17	170.7	0.36	170.4	0.02	225	40
FB -MUL	176.4	0.27	176.8	0.02	3969	10
FC-MUL	173.5	0.31	172.2	0.02	3969	10
FG -MUL	167.5	0.34	172.4	0.02	3969	10
FS-MUL	172.0	0.32	174.0	0.02	3969	10

Table 4.10c Results of pointwise polynomials
ASPECT (degrees)

Convexity statistics are much more difficult to interpret. Profile convexity (Table 4.10d) contains a greater within-surface variance of most parameters. The mean values are all negative, suggesting an excess of concavities in all data sets since, if the convexities balance the concavities, a value of zero is approximated. INCH shows the weakest concavity, while LIAN and FORV show greater concavity. The standard deviations are similarly variable, although INCH and LIAN values suggest a greater degree of concavity consistency. FORV has a larger standard deviation and therefore suggests a greater variation of convexity. INCH has a low negative skewness and low positive kurtosis, suggesting that frequent gentle convexities are balanced by fewer more extreme concavities. LIAN and FORV have similarly low positive skewness and higher positive kurtosis, suggesting frequent high concavities and fewer extreme convexities. In summary, little consistency exists within any surface although across surfaces a limited degree of similarity exists between the data subsets for standard deviation.

Plan convexity (Table 4.10e) is more difficult to express, especially since extreme skewness statistics invalidate the other statistics to a great extent. Again, mean values should be zero if convexity balances concavity, thus while INCH and LIAN illustrate concave surfaces, FORV predominates with convexity. Standard deviations are similarly large for each surface although FORV has the greatest variability. Skewness within each surface shows most inconsistency and is highly variable similarly affecting kurtosis values. Thus, similar to profile convexity, this parameter shows poor surface

AREA	MEAN	S.D.	SKEWNESS	KURTOSIS	MAX	MIN
I49	-1.1	21.9	-0.491	2.791	91	-108
I25	-0.8	10.2	-0.246	1.345	37	-42
I13	-0.4	2.3	0.451	0.365	6	-6
IB -MUL	-1.0	10.0	-0.168	5.914	73	-65
IC -MUL	-0.9	9.7	-0.795	4.800	56	-60
IG-MUL	-0.9	9.9	-0.439	4.979	51	-71
IR-MUL	-1.7	7.2	-0.762	7.265	41	-63
IS -MUL	-0.8	11.1	-0.661	5.321	56	-78
L49	-1.5	26.9	0.641	5.367	226	-115
L25	-1.2	9.5	0.645	6.558	64	-43
L13	-1.3	3.1	-0.178	0.899	9	-11
LB -MUL	-2.0	12.4	1.321	13.139	104	-64
LC -MUL	-2.5	11.1	0.801	10.280	94	-80
LG-MUL	-1.1	7.3	0.725	8.115	62	-47
LR-MUL	-1.4	4.5	0.935	10.227	35	-29
LS-MUL	-1.6	10.7	1.487	12.456	79	-70
F65	-4.1	55.8	1.012	9.199	430	-278
F33	-3.0	22.9	0.407	6.972	124	-142
F17	-1.5	9.9	0.490	3.972	50	-38
FB -MUL	-1.8	20.4	0.606	11.933	176	-130
FC-MUL	-2.0	18.1	1.419	19.838	200	-137
FG -MUL	-1.9	14.5	0.849	17.624	182	-116
FS-MUL	-1.9	15.9	0.689	11.468	124	-145

Table 4.10d Results of pointwise polynomials
PROFILE CONVEXITY (degrees/100 metres)

AREA	MEAN	S.D.	SKEWNESS	KURTOSIS	MAX	MIN
I49	-12.8	254.8	-8.881	212.131	1839	-6728
I25	-13.8	118.2	-9.860	162.501	265	-2080
I13	-1.4	57.0	-1.550	13.256	235	-346
IB -MUL	-13.3	133.1	-10.933	236.625	785	-3461
IC -MUL	-13.7	137.1	-9.641	197.935	1260	-3105
IG-MUL	-10.5	132.6	-12.429	329.373	958	-3832
IR-MUL	-10.7	68.0	-1.075	42.810	937	-678
IS -MUL	-11.1	241.9	-20.245	835.244	3379	-8850
L49	91.3	5361.0	50.141	2564.395	103561	-8616
L25	-0.6	413.7	-3.764	76.715	3342	-6064
L13	-7.3	227.4	4.460	43.013	2016	-654
LB -MUL	-2.6	201.8	6.517	151.414	4800	-2039
LC -MUL	-21.5	236.2	2.139	156.545	5480	-3403
LG-MUL	-11.6	302.4	-10.501	306.522	2463	-8960
LR-MUL	-13.2	222.4	-3.766	83.157	2356	-3791
LS-MUL	-5.0	287.7	-8.776	357.413	4539	-8780
F65	103.9	3665.7	52.741	3130.228	100325	-10536
F33	58.3	730.7	4.242	82.574	11316	-7200
F17	43.8	323.7	4.026	26.822	2765	-741
FB -MUL	2.6	542.8	-4.365	104.970	6499	-11530
FC-MUL	12.0	452.5	-0.038	192.478	8810	-10977
FG -MUL	10.9	448.2	2.243	114.080	9207	-8215
FS-MUL	8.1	416.8	5.330	150.991	11145	-4420

Table 4.10e Results of pointwise polynomials
PLAN CONVEXITY (degrees/100 metres)

consistency with the exclusion of standard deviation, and a weak data set consistency (mean and standard deviation).

Overall, this method of surface characteristic evaluation provides both a good model discriminator and a weak data set discriminator. Altitude provides reasonable success with both model and data sets. The first derivative of altitude - gradient (mean and standard deviation) and aspect (mean) - provide consistent model descriptors - especially aspect. The second derivative of altitude - profile and plan convexity - are weak data set descriptors.

4.4.6 Scale - Variance Components

The scale-variance components were evaluated using the program 'SCAVAL', (Al 4.20), written by the author and based on the theory of Moellering and Tobler (1972). In addition, a small program 'SCAPLOT' was written to generate a graphical summary of the results of SCAVAL (Figure 4.11). The log/log graph shows the effect of different resolutions of grid (y-axis) against the resultant scale-variance components (x-axis). For each surface, the multiquadric interpolated grid of the subsets and the observed grid (REFERENCE) have been plotted. Thus for each data set, a line may be drawn which illustrates the change of variance of the data over a particular range of resolution within that grid. The steeper the line, the more constant the variance is within these resolutions, signifying smoothness and homogenity over that range of sub-area resolutions. As the line flattens, the variance within that resolution range increases, signifying

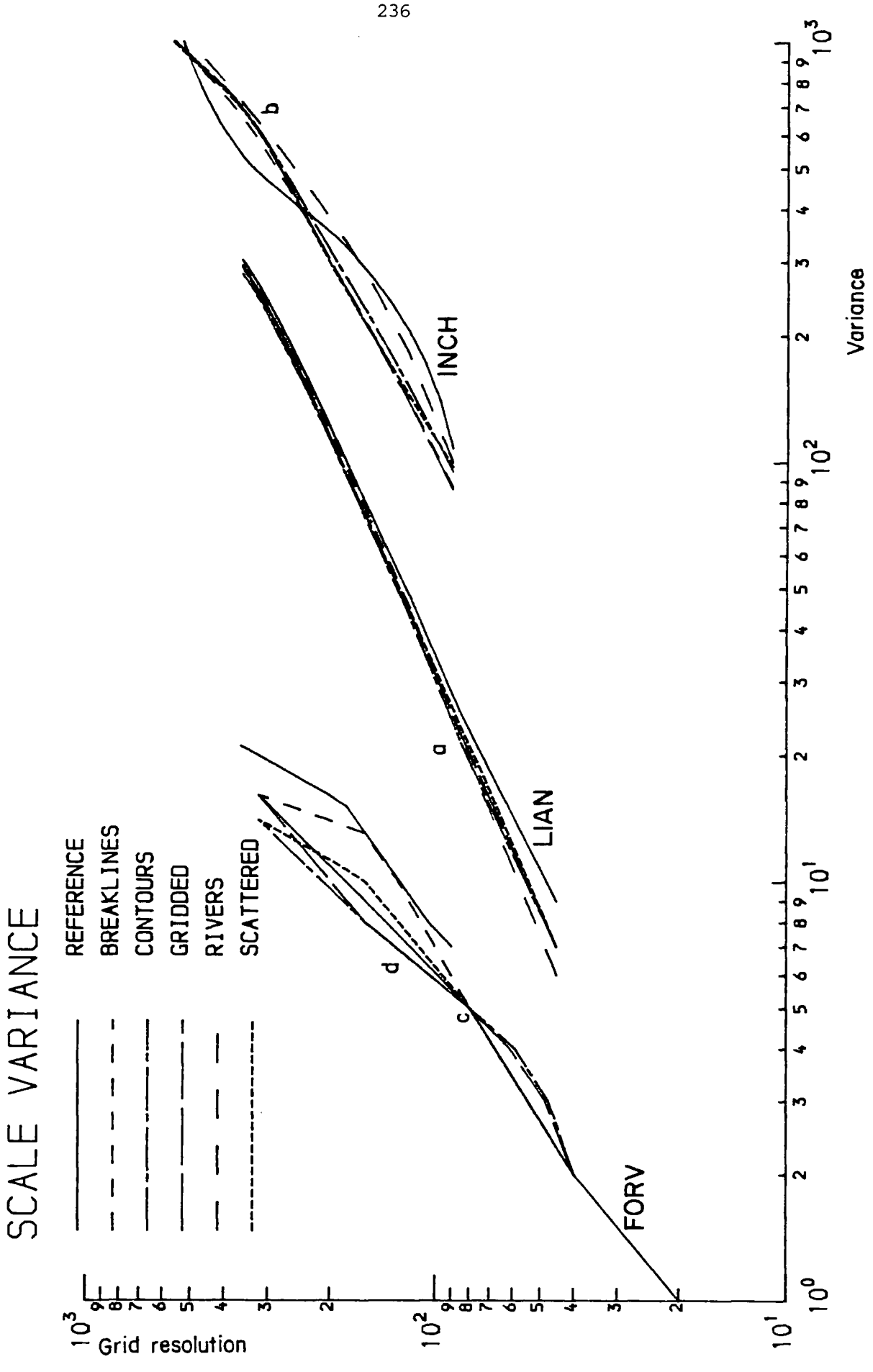


Figure 4.11 Scale-variance results.

heterogenity over the range of sub-area resolutions. In general, a constant gradient (straight line) suggests a constantly changing surface and thus homogenity. Thus, in comparing curves, two aspects must be examined:

a. at the global level - the linearity of the curve;

b. at the global and local level - the gradient of the curve;

On examining Figure 4.11, the three surfaces are quite distinct, thus suggesting the benefits of this method of descriptor for surface identification.

LIAN shows the flattest, most linear function with remarkable consistency between data sets. A very minor break of slope does occur at 90m resolution (a), from which point the lines marginally flatten. This suggests that LIAN is a heterogeneous collection of homogeneous patches, with variance consistent within any patch. The minor break of slope in the curve at 90m suggests that from the next coarser resolution plotted (180m) patches are less homogeneous. This relates to the basic topographic units of the surface - the flat river basin approximately 180m wide and the inclined valley side.

INCH data in general show a similar trend to LIAN, with the exception of the reference grid. The initial variance of the curve at 90m resolution is 100m (compared with LIAN - 20m); thus INCH represents a much rougher surface although this roughness

increases at a similar rate to LIAN. Additionally, a noticeable break of slope occurs at the 360m resolution (b), As the slope of the curve increases after the point, it suggests that more variance occurs beneath the 360/720m resolution than at the more global level. This in turn suggests the opposite of the LIAN situation is occurring i.e. the area may be considered sub-divided into similarly undulating sub-areas of approximately similar mean height.

FORV is more heterogeneous than either of the previous surfaces. While there is more variability between the different data subsets than with either of the previous two surfaces, there is also more variability of variance along any of the curves. However it should be stressed that the variance is minor - of the order of 2m to 15m. Unlike INCH and LIAN, several breaks of slope occur. Over the various data sets, the speed of variance decreases up to 90m (c). This suggests that the surface is highly variable locally, although these are only minor variations on individual topographic units - perhaps small blow-outs on dunes and flat terrain. Above the 160m (d) resolution great conflict occurs between the various data sets. The less evenly distributed data sets (SCATTER and BREAKLINE) show a sharp decrease in increasing variance, suggesting homogeneity of these sub-areas, while the more evenly distributed data sets show an increase in variance suggesting greater heterogeneity at the global level. This is clearly a result of the ability of the data sets to delineate the features of the sub-areas, since the final variances at 320m resolution are remarkably similar.

Unlike the ability of the scale-variance procedure to distinguish surfaces, its ability to distinguish data subsets is very poor. No consistent pattern appears either within a surface, or between the surfaces, mainly as a result of the ability of the subsets to replicate themselves and the reference surface especially in the case of LIAN data.

The INCH reference grid shows a more pronounced difference from the interpolated grids and suggests that, at the local and global levels, multiquadric analysis smooths while it interpolates. In addition, the problems associated with using RIVER subsets are obvious since this increases variance; this is discussed further in Chapter 6.

Many problems exist in analysing the FORV subset/reference variability. Indeed, the reference grid was sampled twice and even this comparison between two reference grids has produced a greater variance than the subset/reference variance. (The finer resolution reference grid curve commences on the y-axis, while the coarser grid is shown by the rightmost curve.) The conclusion must be that the technique of analysis by scale-variance may be too sensitive for such a small range of height, with only subtle differences existing across the DEM as opposed to INCH and LIAN.

In general, then, this technique is an excellent surface discriminator, but very poor data set discriminator.

4.5 CONCLUSION

Overall, six different techniques of describing surface

characteristics or roughness have been examined. The techniques vary dramatically and the data supplied to each may be in point or grid form. An empirical summary of the merits of the various descriptors is given in Table 4.11.

It is apparent that model surface may be differentiated using most of the techniques, although vector analysis, scale-variance and several pointwise polynomial techniques are best. Consideration of surface aspect produces random results since it relates only to the orientation of the model when the data were collected. Altitude is similarly a poor descriptor to choose, since it is scale-dependent, relating to the units of the data. Vector analysis, while successful for similarly-sized grids, suffers in that the descriptors vary with grid resolution. Scale-variance is perhaps the most successful when consistency is taken into account, although two-dimensional autocorrelation and perhaps surface area analysis (regular grid) must be considered as effective descriptors.

Unfortunately, none of the descriptors are completely successful in consistently distinguishing data set type. While this may be a function of using the derived interpolated grids for regular grid descriptors, point-based methods produce variable results. Either some other data acquisition-related factor must exist - perhaps associated to data distribution - or the data set types must be considered to be characteristically different.

These aspects will be developed further in subsequent chapters in an attempt to relate interpolation and isarithm

METHOD	SUCCESS IN UNIQUELY DISTINGUISHING		CONSISTENCY OF	
	SURFACE TYPE	DATA SET TYPE	THIS DESCRIPTOR	SIMILAR DESCRIPTORS
Areal AUTOCORRELATION	**	**	***	**
Two-dimensional AUTOCORRELATION	***	*	***	
Regular-grid SURFACE AREA	***	*	**	
Point SURFACE AREA	**	*		
VECTOR	****		**	
POINTWISE POLY.				
Altitude	****		****	
Gradient	***		**	**
Aspect	****		***	
Prof. Convexity		**		
Plan Convexity				
SCALE VARIANCE	****	*	***	

Table 4.11 Empirical summary of the strengths of various surface characteristic descriptors.

errors to surface characteristics.

CHAPTER 5

GENERATION OF ACCURACY STATISTICS

5 GENERATION OF ACCURACY STATISTICS

5.1 INTRODUCTION

> "Map accuracy is a fundamental issue and of universal concern" (Harley, 1974, 1).

Topographic map accuracy or, more generally, map accuracy is an essential aspect of the mapping art. Marsden (1960) identifies two valid reasons for the utilisation of map accuracy definitions: the existence of private firms and thus contracts demands criteria for the purposes of determining the acceptability of the work and, second, the map user is entitled to have information on the inherent accuracy of the map being used. Additionally, a situation may be envisaged in the development of data banks, where users, by knowledge of the accuracy of the existing data, may select information to match their needs. More immediately, accuracy statistics are of particular relevance in this thesis since the quality of results from interpolation of a surface is an important criterion on which to select the interpolation method.

5.1.1 Factual And Metric Accuracy

Map accuracy contains two components. Factual (or Semantic) accuracy is concerned with planimetric and typographical features. Thus factual accuracy is a measure of the exactness of symbols (point; line; area), colours and names, and is of no concern in this thesis. Alternatively, metric accuracy is

concerned with actual location in space, therefore accuracies in the horizontal and vertical planes must be considered. Horizontal accuracy is only of interest within this thesis in so much as it is generally an element of any vertical accuracy specification.

This chapter will be concerned directly with vertical map accuracy.

5.1.2 Accuracy, Precision, Resolution And Error

It is imperative to define the meanings of 'accuracy'; 'precision'; 'resolution'; and 'error' from the outset. Resolution represents the smallest unit of measurement that may be defined by the measurement procedure. Thus, on a metric scale marked off in millimetres this would be 1mm.

All data contain errors, the magnitude and frequency of which are subsumed under the term accuracy. Three types of error are definable. If the errors are stochastic, then

$$\sum_{i=1}^{n} e_i = 0,$$

where e_i = the error at a point i,

n = the number of data points.

In this case, the Root Mean Square Error (RMSE),

$$RMSE = \left(\left(\sum_{i=1}^{n} e_i^2\right)/n\right)^{1/2},$$

is used to show the consistency of the data. Alternatively, if the error is accumulative or systematic, then,

$$\sum_{i=1}^{n} e_i \neq 0.$$

The mean error is thus given by,

$$\overline{E} = \left(\sum_{i=1}^{n} e_i\right)/n,$$

and the standard error formula,

$$s.e. = \left(\left(\sum_{i=1}^{n} (e_i - \overline{E})^2\right)/n\right)^{1/2},$$

must be used to give the data consistency.

Finally, gross blunders may exist in the data. In topographic mapping, Surveys and Mapping Branch (1977) consider these to be anything over 3 times the standard error.

Precision and, as stated above, accuracy are very closely related to these types of error. If no systematic error exists, then precision will be the same as accuracy. If only systematic error exists, then high precision but low accuracy will be said to exist. Precision is purely a measure of the lack of scatter of data - accuracy is a measure of data bias.

It must be stressed that the map accuracy is the most important item of a map. A map will only be employed when the user has confidence in its portrayal of the 'real world'.

5.2 ASSESSMENT OF HEIGHT ACCURACY

When studying height accuracy or, more specifically, isarithm accuracy, we must distinguish between 'geometric accuracy' and the 'morphological trueness' of the isarithms.

5.2.1 Geometric Accuracy

Geometric accuracy provides an expression of the absolute horizontal and vertical errors of the isarithms or spot heights. This is derived from the map by testing a sample of points which, if sufficiently random, will offer a descriptor representative of the entire area and be well-suited for direct comparison with alternative products.

Generally speaking, elevation error is derived in relation to slope using Koppe's formula. The error at a point, whether on the node of a grid or at a point of interest or data point, can be evaluated easily. Assuming a series of exact or reference values, the error of an observation at the same planimetric position can be found from,

$e = z_r - z_o$,

where z_r = the height of the reference point,

z_o = the height of the observed point.

Alternatively, this elevation error can be translated into a position error by a consideration of the tangent of the surface slope angle at that point,

$e = z_r - z_o = b \cdot t$,

and after transformation becomes,

$b = (z_r - z_o) / t$,

where b = the position error

t = tan (slope).

On the basis of this, Koppe discovered a simple empirical law of errors which can be stated,

$Mh = +/- (e + (b \cdot \tan (\alpha)))$,

where Mh = the mean height error at a point,

e = the RMS height error,

b = a constant horizontal accuracy due to the inherent mapping operations (typically +/-0.2mm at map scale for European maps and +/-0.3mm for US maps),

α = the slope at the point.

Using the Koppe formula, some national map accuracy

specifications are summarised in Figure 5.1. These will be examined further in Chapter 7.

5.2.2 Morphological Trueness

Morphological trueness (or relative isarithm accuracy) represents the ability of the isarithm map to regenerate the configuration of the surface. This is far more difficult to express universally.

One of the main attempts to quantify this error was evaluated by Lindig. Although based on simple theory, the process is very laborious and time-consuming and this has tended to prevent its universal adoption. However, this method should become more attractive in an automated computer environment. Lindig (1956a, 1956b) examined the work of Finsterwalder, Schaefer and Schmidt, and defined four errors visible in a contour. Three of these errors, i.e. position, direction and curvature, can, in principle, be found by direct measurement. The fourth - elevation - is a function of position error and slope. The position and elevation errors are a measure of geometric accuracy, but must be established before the morphological trueness parameters, direction and curvature error, are established. For the purposes of definition, the reference or exact contour (H) will have no superscript and the subject contour that is being compared (H′) will have superscript ′, see Figure 5.2.

a. Position error (dx_i) is the distance, between the two contours, of the line drawn perpendicularly to the tangent of the contour at the point P_i,

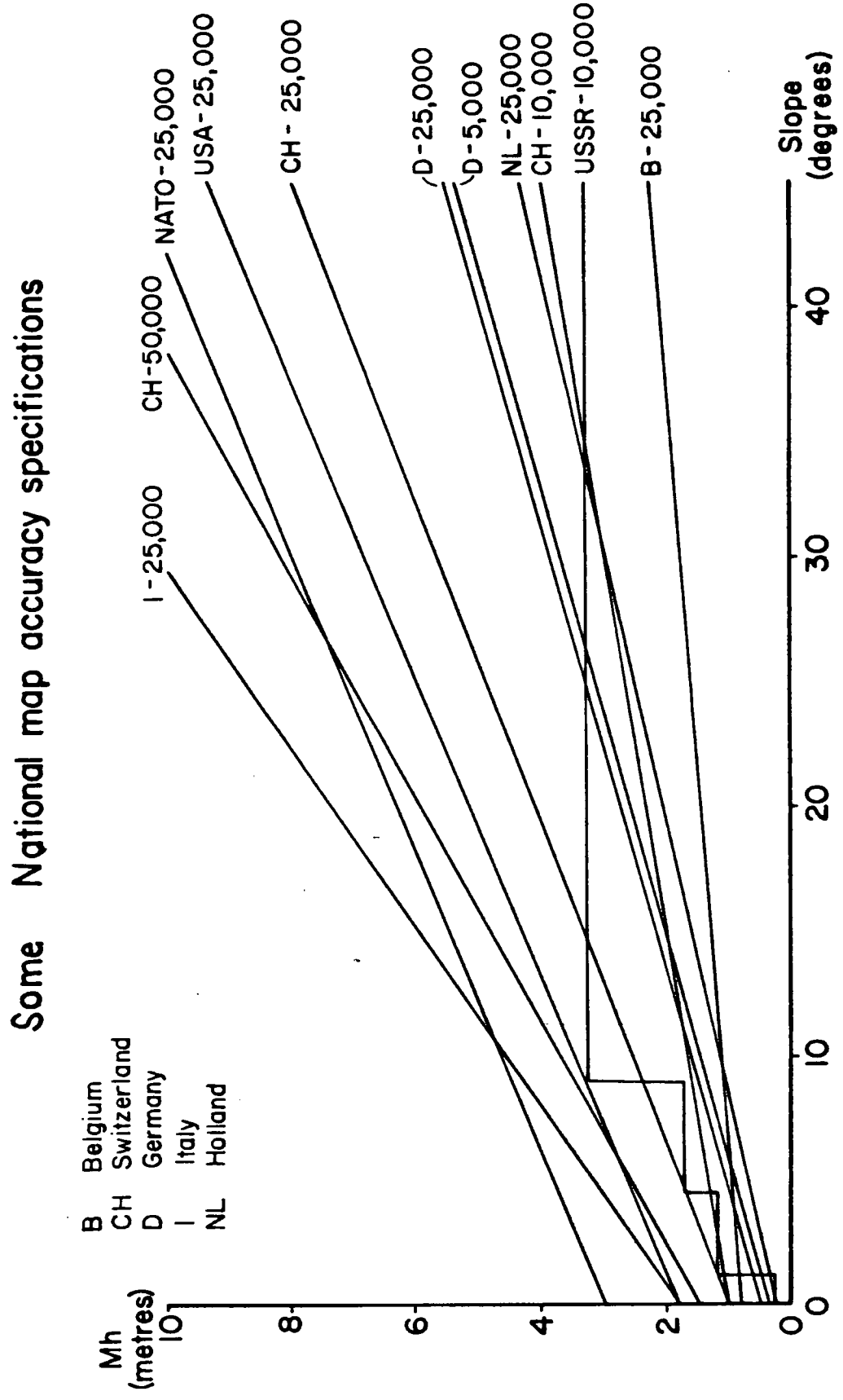
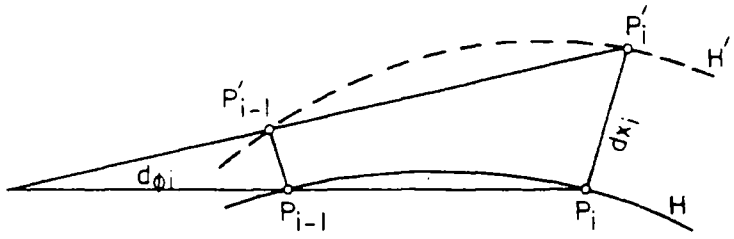


Figure 5.1 Some national map accuracy specifications

Figure 5.2

Morphological trueness (after Lindig.1956)



Position error = $dx_i = P_i - P'_i$

Direction error = $d\phi_i$

Curvature error = $dk_i = k_i - k'_i$

Elevation error = $dh_i = dx_i \cdot \tan(\alpha_i)$

$dx_i = P_i - P_i'$.

b. Direction error ($d\phi_i$) is the angle subtended between the chord P_{i-1} , P_i on the reference contour and the chord P_{i-1}' , P_i' on the subject contour. P_{i-1}' is formed from P_{i-1} in exactly the same way as P_i' from P_i .

c. Curvature error (dk_i) is the difference between the curve k_i of the reference contour at P_i and the curve k_i' of the subject contour at P_i' ,

$dk_i = k_i - k_i'$.

Here the curve may be defined as:

$k_i = R / r_i$,

where r_i is the radius of the curve at P_i and R is a fixed standard radius of curvature which must be established a priori in relation to the data sets being used.

d. Elevation error (dh_i) is essentially the Koppe formula. dh_i is the product of the position error dx_i and the tangent of the terrain slope (α_i) at P_i ,

$dh_i = dx_i . \tan(\alpha_i)$.

The actual measurement of these errors can become difficult. Position error is easily measured directly and thus, with a knowledge of the terrain, the elevation error may be derived. However, the 'form errors' of curvature and direction require some ingenuity to derive. Lindig solved the problem of deriving direction error by the development of a 'Dirometer' - essentially a small rotating protractor mounted on a transparent base.

Three different methods were considered for measuring the curvature error. The 'curvature ladder', a variable arc template, and 'curvometer', a diagram drawn on transparent material, used optical methods of comparison. The 'angle of chord' method employed the direction error angle, measured with the dirometer, in an analytical solution.

5.2.3 Summary

In general, in the assessment of contour accuracy, two errors must be considered, but to varying degrees depending on the map scale and user requirements. At large scales, geometric accuracy is of prime importance, whereas at smaller scales, morphological trueness gains importance. Similarly, contours used to derive engineering volumes must have greater stress on geometric accuracy than must contours derived for orienteering which would require greater morphological trueness.

> "Several specifications may be needed according to the use being made of the map. The photogrammetrist has problems which the land surveyor hasn't and vice versa - similarly research, planning, engineering all have different uses" (Blakney, 1968, 1041).

5.3 COMPARISON TECHNIQUE

> "It has always been generally agreed among topographers that it is practically impossible to prepare a topographic map that is so foolproof that it could not in some way be proven in error by amounts exceeding the accuracy tolerance. Any engineer knows that, if he really wanted to, he could take a topographic map prepared by one of the most reputable topographers and prove that it does not comply with the specified standards of map

> accuracy. Since such an approach is negative, it appears that the testing of a topographic map should be done primarily with the proper attitude. The logical attitude should be to determine whether or not the map has been prepared conscientiously and in accordance with standard practice, and whether or not it is generally in accordance with the specified standards of accuracy.
>
> In testing a topographic map, the test should not be made on only the portions of the map that are impossible or most difficult to compile within the specified standards of map accuracy, such as in steep-sloped, heavily-wooded areas, but rather, it should represent areas that are more typical of the over-all project. The testing procedure should also be done logically" (Alster, 1960, 446-7).

This research is concerned with evaluating the performances of various isarithmic and interpolation programs. In evaluating the accuracy of the interpolated products, it is important to realise the unbiassed nature of this evaluation and consider the various alternative methods available.

The evaluation of the fidelity of data reconstructed from sampled points may be classified under three headings, namely:

a. graphically/empirically- giving a purely visual/aesthetic result which is highly subjective and often difficult to quantify;

b. numerically- giving a very statistical and thus quantifiable result but which may be manipulated by deriving alternative sub-samples of the population based on some criterion;

c. analytically - using some special technique not to

test actual results obtained, but to test the data and methods directly.

In the context of this thesis, the first two methods can be broadened into separate considerations of both the geometrical accuracy of an interpolated grid and the morphological trueness of any interpolated isarithms.

5.3.1 Graphical Evaluation Of Geometric Accuracy

The geometric accuracy of an interpolated grid may be displayed using standard graphical products as discussed in Chapter 2. Most of these products require the prior computation of an isopach grid - a difference matrix between the observed reference grid and the interpolated grid.

The simplest method is to produce an isarithmic map of the isopach grid (e.g. Walters, 1969; Walden, 1972 and Braile, 1978). Although it may be successful where there are substantial constant errors involved, in situations where there are small fluctuations in the error across a map, the resultant isarithmic map is often difficult to interpret. Additionally, where several isopach isarithm maps are being compared and markedly different ranges of error occur, the necessarily different isarithm intervals will create interpretation problems. Isopach isarithms can be usefully employed, but are of less importance in this thesis where large ranges of error may exist in the interpolated products.

An isopach grid has also been used by several workers (e.g. Harrington, 1972 and Franke, 1979) in conjunction with block

diagrams (2.3.2.2). However, this method is perhaps the least successful as the view of the error grid is not planimetrically correct and high errors near to the point of viewing obscure some more distant detail.

In order to simulate relief while overcoming the problems of block diagrams, stereoscopy has been introduced (2.3.2.3). Stereograms of point distributions can easily be generated on the computer (Cubitt and Celenk, 1976). Plotting the reference grid in one colour, (e.g. red), and derived grid in another, (e.g. black) and overlaying these stereograms allows us to compare visually two different grids. However, in practice, due to the small size of the stereogram, the limited resolution of the plotter and the perception of the human eye, only relatively large error/terrain height ranges can be compared. This may be improved by utilising the isopach grid to produce the stereogram.

An alternative method is to utilise some form of profile derived from the grid (2.3.3.1). Kalkani (1977) fitted cubic splines to randomly orientated profiles across a regular grid which are in turn overlain for each grid. There is no doubt that this produces an easily comprehensible result, but while this may be useful in comparing error locally, it is certainly too selective to give an accurate global estimation of the error of the grid. Additionally, the use of cubic splines may unintentionally create unwanted distortion.

5.3.2 Numerical Evaluation Of Geometric Accuracy Of Surfaces

There are several techniques used by many workers to

evaluate numerically the geometric accuracy of interpolated points.

The establishment of a Pearson Product Moment coefficient of correlation is generally preferred to other coefficients of correlation for a regular grid. This technique is used by Merriam and Sneath (1966); Morrison (1970, 1971, 1974a); Walden (1972); and Robinson (1972). Robinson however casts doubt on the validity of relying solely on such a technique.

> "The single coefficient of correlation is little more than a broad generalisation however and it is not very revealing. Comprehension may be increased by mapping residuals" (Robinson, 1962, 414).

Whereas correlation techniques have tended to be used in conjunction with mathematical surfaces, topographers have tended to rely on the derivation of RMSE and the Koppe formula to consider interpolation accuracy. Ebner, Hofmann, Reiss and Steidler (1980) compared a photogrammetrically-derived grid with that interpolated by the 'HIFI' program. Similarily Schilcher (1977) and Ackermann (1978) compared surveyed, photogrammetric, 'SCOP'-interpolated and hand-interpolated contours by utilising 485 common test points from the contours and comparing these with the high accuracy ground surveyed points. Lips (1964) and Assmus and Stanger (1978), while deriving similar end products i.e. RMSE and Koppe, unfortunately derived them by subjective means, both 'arbitrarily' selecting pairs of points from overlain contours. Braile (1978), in his examination of interpolation within aeromagnetic surfaces, removed the edge of the matrix to prevent any 'edge effect' and then calculated the RMSE of

each grid against a hand-interpolated reference. It is unfortunate that the validity of this fine consideration is obscured by comparing computer-interpolated grids with a manually-interpolated grid to generate absolute error. Having derived the RMSE, the logical next step is the derivation of the systematic error and the standard error if systematic error exists. Such statistics have been generated by many workers including Stearns (1968); Shaw and Lynn (1972); Edwards (1972) and Bethel, Crawley, Shepphird and Hussain (1978).

There are numerous other methods preferred by other authors. Morrison (1970, 1971, 1974a) derived a unique ratio - that of the standard error of the residuals against those of the surface values. Shepard (1970), in an evaluation of SYMAP, used similar techniques, dividing the grid into parallel profiles and establishing mean, RMSE and standard error of the reference and residual grids and using these results to establish a chi-square statistic to describe each map.

Merriam and Sneath (1966); Robinson (1972) and Whitten (1975) derived a Percentage total sum of squares(q) also referred to by Davis (1973) as the 'Percentage of goodness of fit'. Similar to a correlation, this is derived from the formula,

$$q = 100 \cdot \left(\frac{(\Sigma z_c^2) - ((\Sigma z_c^2)/n)}{(\Sigma z_o^2) - ((\Sigma z_o^2)/n)} \right)$$

where z_o = the observed data reference,

z_c = the computed data value,

n = the number of data points.

A perfect fit would be 100%.

Perhaps the most interesting technique for comparing the geometric accuracy of contours, is the 'Resemblance Matrix', used by Court (1970) with its associated 'medial correlation' coefficient 'q'. While this has not been used previously to derive the geometric accuracy of contour maps, the labour-intensive method is a potentially valid one in a fully automated environment. Court uses it initially to compare quantile isopleths on socio-political mapping.

Assume two binary distributions (see Figure 5.3). If two choropleth maps of these distributions are overlain, then four regions may be identified: where both are unity, (1,1); where both are zero, (0,0); where the first is unity, the second zero, (1,0); and vice versa, (0,1). Replacing the boundaries in the choropleth maps with contour lines, 1 becomes '+' and '0' becomes '-' (Figure 5.3a). The Resemblance matrix may now be established from the sum of the percentage areas within each sub-area (Figure 5.3b). q, the coefficient of medial correlation, is derived from the formula,

$$q = (((++)+(--))-((+-)+(-+)))/((++)+(--)+(+-)+(-+)).$$

Thus a value for the geometric accuracy of a contour map may be established by successively comparing each generated contour with its associated reference contour and meaning the various coefficients for each map.

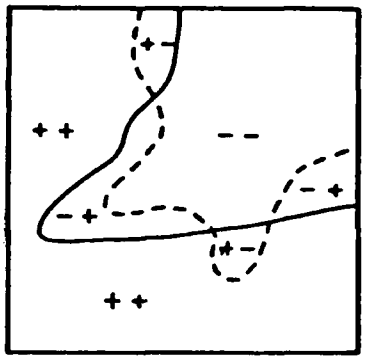
5.3.3 Graphical Comparison Of Contour Morphological Trueness

A graphical comparison of the morphological trueness of

FIGURE 5.3

Contour comparison utilising a Resemblance Matrix

(a) Contour map



——— Reference

- - - - - Interpolated

(b) Resemblance Matrix

Reference \ Interpolated	+	−	Σ
+	50	5	55
−	15	30	45
Σ	65	35	100

$$q = \frac{(50+30) - (15+5)}{100}$$

contours is used by most workers and is generally backed up in the more serious papers by some numerical figures to give some quantifiable geometric error estimation.

In the worst cases, the comparison takes the form of two distinct adjacent, or even separated, plots neither of which is an observed reference (e.g. Whitten, 1970; Rhind, 1971 and 1972; Connelly, 1971 and Harrington, 1976). These maps are obviously difficult to compare as they are not in register or overlain and also only give a relative comparison.

For a more absolute solution, various authors have manually hand-contoured their own data and used this as a reference (e.g. Walters, 1969; Swindel and van Andel, 1969; Walden, 1972; Chiles and Chauvet, 1975 and Braile, 1978). A more rigorous solution is only available when the nature of the data allows direct contouring (e.g. photogrammetric) and contours to be compared can be laid adjacent to directly generated reference contours (Ebner, Hofmann, Reiss and Steidler, 1980). However, in all these cases problems still occur due to the contours not being overlain.

For an accurate solution the contour maps must be overlain (e.g. Zarzycki, 1976; Ackermann, 1978 and Bethel, Crawley, Shepphird and Hussain, 1978). In order to avoid confusion the areas between the same contours may be shaded. Royal College of Art (1972) used this method to compare hand-interpolated bathymetric and geochemical contour maps and digitised O.S. topographic contours with those generated by the SACM and ECU packages.

Where an absolute morphological comparison is required, a lack of sufficiently accurate reference contours has meant a devaluation of the results. In most data surfaces there may be limitations to obtaining contours, in which case workers should avoid these in preference for complete data sets which include directly generated contours, if any form of accuracy statement of interpolation methods is to be established. This restricts data sets to those as discussed in 3.3.1.

5.3.4 Numerical Evaluation Of Contour Morphological Trueness

As discussed earlier, it is the morphological trueness of contours which is the most difficult aspect to quantify numerically. However, the widespread use of computers, and the fact that contours are now frequently digitised, have meant that at least two methods are now easily developed.

Feder and Freeman (1966) have developed a method based on 'pattern recognition' which allows contours to be correlated numerically. Basically, it involves assigning an integer value for any given direction of a line segment (Figure 5.4). Thus the line is numerically directionally digitised (quantised) by superimposing a mesh on the line. The line or 'chain', now 'chain encoded', is in a suitable form for correlating with another chain.

Given two chains R (segment) and S (whole curve),

$R = r_1, r_2, r_3, \ldots, r_n$, and

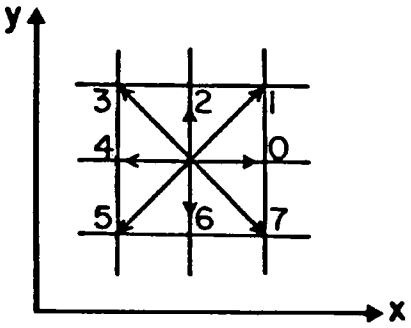
$S = s_1, s_2, s_3, \ldots, s_n$; where $n \leqslant m$,

then

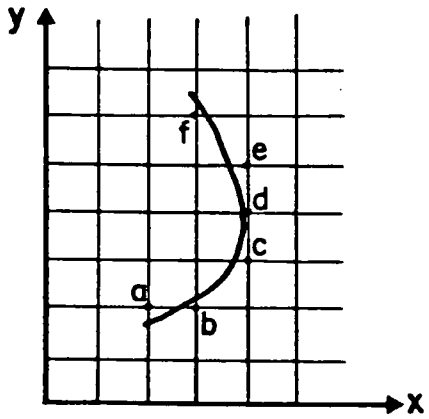
$$q_{ab}(j) = 1/n \sum_{i=1}^{n} a_i b_{i+j},$$

FIGURE 5.4

Contour comparison by Chain quantization
(after Feder and Freeman, 1966)



(a) Octal Encoding Scheme for each line segment



(b) Quantization Scheme

where $a_i b_{i+j} = \cos(\angle a_i - \angle b_{i+j})$,

$-1 \leqslant q_{ab}(j) \leqslant 1$

j = each position of the segment relative to the whole curve (Feder and Feeman, 1966).

The number of directions and resolution can be determined by the user to give a more sensitive descriptor.

Another method, first discussed in 1956, derives the position, direction, curvature, and elevation errors defined by Lindig (1956a and 1956b). As mentioned earlier, this was originally labour-intensive and impractical. However, it lends itself to automation and must be considered as the main basis for contour morphological comparisons.

5.3.5 Analytical Methods

Since the early seventies much interest has been shown within the ITC in the Netherlands in purely analytical methods of establishing accuracy of interpolation methods in digital terrain modelling. Several papers by Makarovic (1972,1974); Clerici and Kubik (1974); Botman and Kubik (1979) and Tempfli and Makarovic (1979) discuss the evaluation of transfer functions and their application to the examination of data reconstruction.

The concept of their analytical approach is based on the premiss that input data can be represented by a Fourier Series. The effect of sampling on the fidelity of the reconstructed data can therefore be studied utilising sine waves. Within this approach, the effects of sampling density

and measuring errors can be evaluated separately and combined by mutual multiplication of the corresponding transfer ratios.

There is no doubt that this attitude to interpolation comparison is a valid one and should be recognised. This thesis, however, is substantially concerned with the empirical evaluation of the accuracy of certain packages, data surfaces and data sets and therefore this analytical approach will not be considered further.

5.4 COMPARISON PROGRAMS DEVELOPED

With reference to contour accuracy estimation, Imhof suggests that,

> "perfect evaluation techniques will probably continue to elude us. We are trying to compare an inadequately defined irregular surface with an incomplete line map with large unmeasurable gaps - based on linear or quadratic change" (Imhof, 1965).

Within this research, this evaluation had to be made and thus various requirements were established for any method of accuracy derivation.

a. The techniques had to derive both geometric and morphological accuracy and be capable of being applied to a range of data at any scale.

b. The techniques had to allow for a 'quick-look' comparison as well as a precise estimation of the accuracy. Various interpolation techniques were being used with many data sets generating numerous results.

c. Additionally, the output from any comparison had to be compact, thereby restricting it to more global comparisons.

d. The procedures must allow for isarithm and/or regular grid information.

On the basis of these requirements, and considering the literature, a strategy using four methods was employed.

a. For each interpolated grid,

a.1. a simple overlain isarithm plot was generated comparing the interpolated grid surface with the reference isarithms. This would provide a 'quick-look' product.

a.2. an in-depth statistical evaluation was performed using 'MATANN' (developed by the author - see A1.3.14 and Chapter 6).

b. For the interpolated isarithms,

b.1. a global expression of their morphological trueness in relation to the photogrammetrically—derived

contours was evaluated by the program 'CONLIN' (developed by the author - see A1.4.5 and utilising the theory of Lindig - see 5.2.2).

b.2. a global expression of their geometric accuracy was similarly established from the program 'CONCOR' (developed by the author - see A1.4.4 and utilising the theory of Court - see 5.3.2).

b.3. a 'quick-look' isarithm overlay was also established.

Additionally, through CONLIN, the Koppe formula for each isarithm map was established to evaluate the ability of the computer to generate isarithms within the tolerance of some national mapping organisations. Perhaps this comparison may be unfair, since in a non-automated system,

> "contours are the result of the manual work of cartographers, and therefore the cartographers' personal interpretation of the relief of the surface. The user has no idea of the accuracy of the contour system itself" (Yoeli, 1966, 254).

CHAPTER 6

RANDOM-TO-GRID INTERPOLATION

6 RANDOM-TO-GRID INTERPOLATION

6.1 INTRODUCTION

It has been demonstrated in previous chapters that random-to-grid interpolation is the kernel of all non-triangulation computer contouring of random data. It is therefore important that this aspect be investigated thoroughly by evaluating the accuracy of individual methods of interpolation and by considering the relationship between interpolation error and data and surface characteristics.

6.1.1 Résumé Of Previous Research

Within the literature there are several examples of the empirical evaluation of random-to-grid interpolation methods. Invariably these are concerned with promoting a certain method of interpolation rather than conducting a thorough research (for example Braile, 1978 and much of the work on kriging), or involve a relatively superficial comparison of the general properties involved in the various interpolation schemes (Rhind, 1971, 1975) or consider only specific situations (Davis, 1975).

Few attempts have been made to perform a more universal, detailed investigation on computer-interpolated surfaces, notable exceptions being the work of Morrison (1971 and 1974b), and Walden (1972).

Morrison has performed the most detailed investigation in considering six independent variables - surface variation, sample size, sample type (or scatter), the interpolation model, directional constraints with the model and size of interpolated grid. From the interpolation results, he computed two statistics - an interpolation/reference grid correlation coefficient and the ratio of the standard deviations of the interpolated values to the standard deviation of the known surface. ANOVA tests were later performed on these statistics to relate them to the previously defined variables. Unfortunately, he only considered inherently smooth, highly correlated mathematical surfaces, from which he concluded that:

a. a minimum of 49 points and a maximum of 100 points are required to perform random-to-grid interpolation;

b. no data set with a NNS (Nearest Neighbour Statistic) less than 0.9 should be used for interpolation;

c. any data set with a NNS greater than 1.25 will capture the statistical properties of the distribution form accurately;

d. point and areal functions perform equally well, although weighted and planar functions are best-behaved;

e. in general, weighted averages should be used with data sets of less than 100 points, planar for approximately 100 points, and quadratic surfaces for greater than 100 points;

f. the more complex the surface, the larger the error, and the less the effect the point scatter has.

It is unfortunate that Morrison did not relate his otherwise precise statistics to reality. As Rhind suggested

> "the trade-off factor between different point patterns is totally unknown - would 5000 data points with a nearest neighbour value of 0.8 produce as reliable an end result as 1000 points with a corresponding value of, say, 1.3?" (Rhind, 1971, 155).

Additionally, is such a global (all data) technique meaningful for complex surfaces? If the NNS is 0.8, should the data be filtered to increase the NNS to, say 1.0 ? - this seems unlikely.

Walden (1972) performed a more specific investigation, although she only considered four of Morrison's variables - surface variation, sample size, the interpolation model and the size of the interpolated grid. The surfaces used were more variable in type, a mathematical function (sphere), sub-surface geological data and topographic data. Unfortunately, since the reference grids for the topographic and geological data were generated by interpolation utilising

SURFACE II GRAPHICS (Sampson, 1978), the later results were biassed to similar interpolation techniques. Sample sizes chosen were 100, 250 and 400 points and these were interpolated using a selection of mainly Kansas Geological Survey published algorithms and commercial packages, producing three basic grid sizes (25x25, 41x41 and 61x61). In addition to creating standard error statistics for each data/interpolation/grid size combination, Walden attempted to correlate error with the distance from the grid node being interpolated to its nearest neighbour and the number of data points within a fixed range of the grid node being interpolated.

Walden (1972, 93-101) concluded that:

a. "In general, increases in sample size are directly related to increases in the reliability of the entire map", although global fit methods have less improvement. This appears to negate Morrison's 100 point theory.

b. "Except for grids generated by SYMAP, the accuracy of surface estimation at the grid nodes seems relatively unaffected by changes in grid size." SYMAP accuracy is related to grid size since co-ordinates of a data point are assumed to be the centre of the cell (or pixel) which contains the point.

c. "The relationship between estimation error at a grid

node and its distance to the nearest data point is not as high as had been expected." In some respects, this is not surprising as it represents a very crude measure.

d. "Consideration should be given to the degree of the surface complexity in the selection of sample size and in selecting the gridding algorithm to be used." Walden did not however propose any relationship.

6.1.2 Evaluation Procedure

Within this research, fourteen data sets, representing three surfaces (see 3.3), were processed using fourteen different interpolation methods (see 2.5). While great care was taken with all programs to ensure that sensible parameters were input, standard default values were used where possible. It is conceivable that improved results could be achieved by experienced users of certain programs. However, it is impossible to examine all parameter combinations and certainly the average user, not least in academia, must be considered inexperienced. The results of this empirical examination will be examined primarily according to the range of the interpolating function (i.e. pointwise, patchwise, global and hybrid) and, secondly, arranged according to each interpolation method/program.

For the purposes of this examination, various statistics were derived for each random-to-grid interpolation performed.

In each case, the interpolated grid was compared with the relevant observed reference grid using the program 'MATANN' (A1.4.13). Initially, an error matrix was generated which formed the basis of all the error statistics.

It was essential to perform a preliminary investigation of a cross-section of these error distributions so that the results could have a bearing on the precise form of the final statistics used. This was undertaken on a full range of the data surfaces, data set types and interpolation methods. From this study, it was concluded that:

a. The initial error distributions generated were near-normal; skewness was generally low and positive, although kurtosis was often more peaked than normal.

b. In many cases, particularly when poorly distributed data set types were considered or certain interpolators were used, several extreme errors were generated. In such circumstances, these outliers completely distorted and invalidated certain statistics - e.g. mean, standard error and kurtosis.

c. In addition to purely error-based statistics, the cross-correlations between absolute error and various neighbour- and gradient-based distributions were investigated. It is undesirable that any distribution being correlated with another should be skewed, therefore various

methods of distribution transformation had to be examined.

As a result of these conclusions, 34 statistics were computed for each data set/interpolation method combination as follows.

a. %BAD. This statistic gives a descriptor of the consistency of the interpolation. It is generated by computing the number of grid nodes (as a percentage of the total) where the absolute error is greater than an arbitrary value (40m). This value was selected since, on the basis of the preliminary investigation, it was considered that this would only indicate problem areas where the interpolator had clearly failed in relation to the input data. It would also allow a method whereby any errors flagged by an interpolation program (e.g. SURFACE II GRAPHICS), which could not compute an individual grid value, could be noted.

b. ORIGINAL ERROR RANGE. Statistics which provide more information on the consistency of the interpolation are the minimum and maximum values of the original error distribution (theoretically, more correctly termed the maximum negative and maximum positive error). These values are derived from the original (interpolation - reference) difference surface and therefore should be of near similar magnitude if a normal distribution exists. Any deviation from equality should indicate a particular bias in an interpolation process, although it should be stressed that

this is a point-based statistic and therefore more responsive to sharp local fluctuations. Also, where interpolation programs could not compute a value and have flagged the resultant interpolation grid, extreme errors have been ignored, thus biassing their statistics. In such cases the tabulated values have been enclosed in brackets.

c. TRIMMED ERROR RANGE. In order to overcome the problems of coping with extreme errors as discussed earlier, a quantile-based approach was used for the subsequent statistics. Having considered the %BAD statistics, arbitrary values of 2.5% to 97.5% (equivalent to a 95% confidence limit - or 2 standard deviations of a normal distribution) were used as the lower and upper quantile levels. These limits were selected as it was considered they removed all gross blunders and extreme errors (greater than 2 standard deviations), while not unnecessarily biassing subsequent statistics. The resultant quantile-based minimum and maximum errors (also more correctly termed the maximum negative and maximum positive error) therefore give a more stable descriptor of the range of the acceptable interpolation errors. Unfortunately, in situations (eg. with the use of SURFACE II GRAPHICS on certain data sets) where there were more nodes not interpolated than nodes removed by trimming, the sample size had to be reduced even further, since clearly these values could not be included.

d. MEAN ERROR. The arithmetic mean of the nodal errors represents the systematic error involved in a particular interpolation. The mean is used here in sensu lato and is strictly speaking a media di conto. It is the commonest measure of central tendency which is readily understood, easily computed, and is a useful benchmark for comparison with other evaluations external to this research. Its main disadvantage, that its value is distorted by one or two extreme values, has been overcome by its derivation from the quantile-trimmed error distributions.

e. MEDIAN ERROR. It is expected that a measure of central tendency should be near the middle of the distribution to which it refers. The median, whose value divides the total frequency into two halves and therefore is not changed by trimming, must therefore be considered a superior statistical measure to the mean. It has the advantage over the mean that it is not biassed by the rest of the distrbution, and therefore illustrates whether positive or negative errors are more likely. However, this could be argued to be a disadvantage in specific circumstances. In general, the strength of using both measures lies in their mutual complementarity.

f. MEAN OF THE ABSOLUTE ERROR (ABS MEAN). This is at least as important as the MEAN ERROR, since it gives an estimator of the magnitude of the likely error of the

process. This mean is computed in exactly the same way as the MEAN ERROR except that, after the quantile trimming, the absolute error has been used. By definition, it has the disadvantage that the distribution is truncated at 0.0, and therefore there is a high probability that it is positively skewed. This would be unsatisfactory if the statistic was to be used in subsequent statistical analyses, as it could be argued that it was not truly representative of the central tendency of the distribution.

g. MEDIAN OF THE ABSOLUTE ERROR (ABS MEDIAN). This also gives a superior measure of central tendency. This is especially true since the absolute error distribution is usually positively skewed. It therefore provides a more statistically satisfactory measure if a statistical measure of central tendency is to be used in a subsequent statistical analysis. As previously, the strength of using both median and mean measures lies in their mutual complementarity. This median is computed in exactly the same way as the MEDIAN ERROR except that, after trimming, the absolute error has been used.

h. STANDARD ERROR. This provides an estimator of the variance of the error distribution. It provides an indicator of how error is distributed about the mean, and has therefore frequently been used as a standard measure of the global accuracy of the interpolation by other

researchers. Since it is the second moment of a frequency distribution taken about the arithmetic mean, it suffers from the limitations of the mean. This makes it susceptible to extremes and skewness. It was therefore computed from the trimmed distributions to remove the effects of extreme values. However, since the measure is frequently used in similar studies, the distribution was not transformed to minimise skewness (although this may have been preferable for any subsequent statistical analyses), so as to enable other researchers to perform direct comparisons with their statistical results.

i. SKEWNESS and KURTOSIS. These were computed of the full quantile-truncated error distributions to present a more concise descriptor of the shape of the distributions. SKEWNESS measures the asymmetry of a distribution. In a Gaussian distribution the mean, the median and mode coincide. In a unimodal skew distribution, the mean and - to a lesser extent - the median lie on the side of the distribution which has the longer tail although the mode is still at the point where the curve is highest. Positively skewed distributions have the longer tail to the right. Caswell (1982) suggests that while the value of the coefficient of skewness lies between -3 and +3, values below -1 and above +1 are rare and indicate very skewed distributions.

KURTOSIS is

> "used to describe the extent to which a unimodal frequency curve is 'peaked'; that is to say, the extent of the relative steepness of the accent in the neighbourhood of the mode" (Kendall and Buckland, 1982).

In this situation it assumes a value of 3 for a normal distribution. Curves for which the value is less than, equal to or greater than 3 are known respectively as platykurtic, mesokurtic and leptokurtic.

j. SURFACE CORRELATION. The SURFACE CORRELATION (that correlation between the interpolated and reference grid nodes) is another popular descriptor which conveys an estimate of the global accuracy of the interpolation. Unfortunately, due to the greater within-surface variances and smaller covariances, virtually all interpolation surfaces have SURFACE CORRELATIONS of greater than 0.950 with the majority greater than 0.987. This devalues the statistic's comparative descriptive power. It should be noted that, in computing this and the subsequent Pearson's product moment coefficients of correlation, grid nodes with errors outside the quantile range were not included.

k. ERROR/CHARACTERISTIC CORRELATIONS. In attempting to relate absolute error to characteristics of the data, three correlation coefficients were computed for each interpolated grid - ERROR/NEAR, ERROR/NONN and ERROR/GRAD; it should be stressed from the outset that their absolute magnitude is not considered nearly as important as the

trends they display. In computing any correlations, it is important that, where possible, the two contributing distributions should be as near-normal as possible with low skewness. It was therefore decided to transform all component distributions to minimise their skewness.

After a detailed examination of possible transformations and their effects on a representative range of the distributions to be encountered. three solutions were selected, namely:

1. no transformation;
2. $\log_{10}(x)$ transformation;
3. $\log_{10}^{2}(x)$ transformation.

Since virtually all distributions were strongly positively skewed it was hoped that the log transformation would reverse this skewness. In practice, the reversal was too strong and a moderate positive influence was generated by squaring the log transformation. It is impossible to consider the effect on all the specific situations in this limited discussion; however, since the GPCP interpolation method produced approximately average results, it has been used as a specific case study - see Tables 6.1a-d. The four tables summarise the SKEWNESS and KURTOSIS statistics for each solution. For each data set, the superior solution is enclosed within a box. While no one solution is optimum, the $\log_{10}^{2}(x)$ transformation is generally superior and is

	TRANSFORMATION					
	none		$\log_{10}(x)$		$\log_{10}^{2}(x)$	
DATA	SKEW	KURT	SKEW	KURT	SKEW	KURT
IB	1.29	4.17	-0.47	2.64	0.28	2.24
IC	0.93	3.12	-0.70	2.97	0.02	2.11
IG	0.94	3.18	-0.86	2.69	-0.01	1.96
IR	1.62	5.20	-0.17	2.35	0.60	2.48
IS	1.09	3.66	-0.65	2.93	0.12	2.17
LB	1.16	3.65	-0.55	2.75	0.20	2.15
LC	0.82	3.02	-0.91	3.52	-0.14	2.23
LG	1.05	3.58	-0.93	3.04	0.00	2.12
LR	1.34	4.16	-0.41	2.55	0.35	2.25
LS	0.94	3.21	-0.71	3.05	0.02	2.11
FB	1.90	6.70	-0.21	2.49	0.68	2.85
FC	1.18	3.89	-0.59	2.85	0.18	2.22
FG	1.45	4.85	-0.63	2.74	0.33	2.32
FS	1.73	5.92	-0.33	2.58	0.56	2.61

Table 6.1a Summary of SKEWNESS and KURTOSIS after transformation of the ABSOLUTE ERROR distribution.

	TRANSFORMATION					
	none		$\log_{10}(x)$		$\log_{10}^{2}(x)$	
DATA	SKEW	KURT	SKEW	KURT	SKEW	KURT
IB	1.09	4.10	-0.57	3.52	0.08	2.61
IC	1.12	5.99	-0.90	4.55	-0.14	3.10
IG	-1.54	4.59	-2.39	6.86	-2.22	6.38
IR	1.38	5.40	-0.54	3.27	0.20	2.69
IS	0.76	3.55	-0.87	4.35	-0.20	2.76
LB	1.52	5.33	-0.31	3.01	0.42	2.68
LC	2.03	7.54	-0.07	3.60	0.77	3.64
LG	-0.19	3.44	-2.62	8.56	-1.83	6.30
LR	1.47	4.74	-0.27	2.71	0.45	2.52
LS	1.35	5.88	-0.58	3.63	0.12	2.92
FB	1.54	5.98	-0.35	2.87	0.33	2.92
FC	0.78	3.39	-0.86	4.07	-0.18	2.65
FG	-0.26	2.90	-2.89	10.92	-1.66	6.21
FS	0.88	3.51	-1.29	6.28	-0.12	2.70

Table 6.1b Summary of SKEWNESS and KURTOSIS after transformation of the NEAR distribution.

	TRANSFORMATION					
	none		$\log_{10}(x)$		$\log_{10}^{2}(x)$	
DATA	SKEW	KURT	SKEW	KURT	SKEW	KURT
IB	0.58	2.67	-1.43	4.66	-0.33	2.29
IC	0.41	3.31	-0.90	4.55	-0.14	3.10
IG	1.96	6.05	1.59	3.45	1.56	3.71
IR	1.18	4.30	-0.58	1.98	0.19	1.96
IS	0.71	3.53	-1.90	8.02	-0.40	3.02
LB	1.75	7.14	-0.42	1.96	0.50	2.40
LC	0.25	2.29	-1.61	4.65	-0.67	2.61
LG	-0.34	2.33	-2.80	17.94	-1.04	4.04
LR	1.09	3.27	-0.07	1.36	0.43	1.71
LS	0.72	3.30	-1.64	5.33	-0.42	2.72
FB	0.64	2.87	-1.22	3.53	-0.30	2.13
FC	0.91	3.88	-1.66	7.67	-0.23	2.94
FG	0.67	2.16	-1.07	11.16	0.09	2.80
FS	0.61	2.70	-1.47	5.14	-0.30	2.37

Table 6.1c Summary of SKEWNESS and KURTOSIS after transformation of the NONN distribution.

	TRANSFORMATION					
	none		$\log_{10}(x)$		$\log_{10}^{2}(x)$	
DATA	SKEW	KURT	SKEW	KURT	SKEW	KURT
IB	1.33	5.80	-0.75	4.07	0.07	2.98
IC	1.27	5.62	-0.78	4.14	0.03	2.99
IG	1.18	5.19	-0.81	3.95	-0.02	2.87
IR	1.27	5.62	-0.78	4.14	0.03	2.99
IS	1.30	5.92	-0.81	4.03	0.00	2.95
LB	0.74	2.97	-0.55	2.71	-0.02	2.01
LC	0.72	2.93	-0.57	2.72	-0.04	2.00
LG	0.74	2.84	-0.50	2.65	0.02	1.97
LR	0.80	3.04	-0.48	2.61	0.05	1.99
LS	0.77	3.01	-0.50	2.65	0.02	1.99
FB	2.63	13.61	0.01	2.79	0.97	4.10
FC	2.59	13.20	-0.03	2.87	0.91	3.99
FG	2.64	14.96	-0.03	2.78	0.91	4.07
FS	2.66	14.45	-0.01	2.81	0.95	4.11

Table 6.1d Summary of SKEWNESS and KURTOSIS after transformation of the GRAD distribution.

never the worst solution. It was therefore decided that all the relevant distributions would undergo this transformation before any correlation coefficient was generated. The poor SKEWNESS values for the LIAN data sets in the NEAR and NONN distributions, the FORV data sets in the GRAD distributions and the GRID data sets in the NEAR distributions should be noted for the main discussion on the resultant correlation coefficients.

Another factor to be considered is the level of significance for the correlation coefficients. These have been computed for three levels of significance for each of the data models and are summarised in Table 6.2. These are remarkably low as a direct result of the very large populations used to generate them - 2400-4000 depending on the specific data models used.

More specifically, the ERROR/NEAR coefficient represents the correlation between the absolute trimmed error and the distance to the nearest data point (at the relevant grid node) used to generate that grid. If the accuracy of the interpolation varies directly with the nearness of a data point, then this value should tend to +1.0.

The ERROR/NONN coefficient represents the correlation between the absolute trimmed error and the number of data

SIGNIFICANCE	INCH	LIAN	FORV
very highly significant (0.1%)	0.067	0.061	0.050
highly significant (1%)	0.053	0.048	0.041
significant (5%)	0.040	0.036	0.030

Table 6.2 Levels of significance for interpolation/ characteristic correlation coefficients.

points (or NO. of Nearest Neighbours) within a certain range of the relevant grid nodes. The ranges used were constant for each data model: 95.0m for all INCH data sets and 50.0 for all LIAN and FORV data sets (the two different values being necessitated by the major difference in scale and area covered by INCH with respect to LIAN and FORV). Since interpolation accuracy is generally thought to improve with the average number of data points used to establish each interpolated point, a negative coefficient would be expected. However, it must be remembered that the NONN distribution is discontinuous and this may adversely affect the correlation coefficient.

The ERROR/GRAD coefficient is the correlation between the absolute trimmed error and the gradient (at the relevant grid node) of the reference terrain surface. In conventional topographic mapping, vertical error is always related directly to gradient. Therefore, as the gradient increases, so the error should increase, resulting in a positive correlation.

1. IN/OUT RANGES and MEANS. Two final sets of statistics were generated purely to aid the investigation of the relative causes of interpolation error between specific data sets and interpolation models. They were therefore selected to present an easily understood descriptor which could be easily compared between data sets/interpolation

models. The IN-MIN, -MAX and -MEAN values follow on from the previous correlation coefficients in that they provide more information on the NEAR, NONN and GRAD values which occurred at the trimmed error grid nodes. Alternatively, the similar format OUT-MIN, -MAX and -MEAN versions are derived from the NEAR, NONN and GRAD values which occurred at the error grid nodes outside the quantile limits. They therefore provide indicators, when compared with the IN statistics, as to the likely causes of extreme error.

These statistics are secondary compared with the previous statistics which quantify the interpolation error, however, they do aid the investigation of error sources.

In conclusion, a wide variety of statistics have been generated for each data set/interpolation method combination. These have been arranged in two summary tables for each interpolation method - the second table containing the secondary IN/OUT statistics only. In addition, a series of rank correlations have been established between these statistics and those data and surface characteristics generated earlier. These have been generated within each model, and averaged overall. From the previous discussion in Chapter 4, it was apparent that LIAN-RIVER produced unrepresentative results due to unusual features of the data. That data set was therefore omitted in the bulk of the computations of the rank correlations. It was felt necessary to illustrate the strength of the correlations by using scattergrams, however, it was also felt that too many

scattergrams would unnecessarily increase the thickness of this thesis, detract from the importance of the other statistics, and overstress the only relative importance of the correlations. Therefore, sample scattergrams have only been included in the conclusion to this chapter (6.6.3 and 6.6.4).

Throughout the discussion, several examples of the isarithm maps derived from the interpolated grid (using GHOST-CONTRA) are used to illustrate the general morphological 'trueness' of the methods of interpolation. Clearly, it is impossible to examine all the data sets in such a fashion, as this would result in several hundred maps. A representative subset of the full total has therefore been carefully selected to illustrate the nuances produced by the various methods. Within these maps, isarithms were drawn both at the same levels as the reference isarithms and, where necessary, at two levels below and above these levels. Therefore in several cases (notably GINOSURF, GHOT and MINCURV), where extreme interpolated grid values occured outside this range, no isarithms were drawn: in many of these cases, their inclusion would have resulted in intolerably long plotting times and the dense isarithms would have obscured all other details in their vicinity.

It should be noted that the titles of the figures have been abbreviated. The first letter refers to the data model (INCH - I, LIAN - L and FORV - F). The second letter refers to the data set type (BREAKLINE - B, CONTOUR - C, GRID - G, RIVER - R and SCATTER - S). Thus, 'F-S OCTANT' refers to the results of processing the FORV-SCATTER data set through the OCTANT option of SURFACE II GRAPHICS. In addition, the standard convention

throughout this discussion will be to refer to the top figure as Figure a and the bottom figure as Figure b.

6.2 POINTWISE METHODS

6.2.1 SURFACE II GRAPHICS

Grid estimation may be performed in SURFACE II GRAPHICS by utilising three types of function (see also 2.5.3 and A1.3.12). The 'TREND' routine involving global polynomial interpolation was, however, not available in the version accessed. 'KRIGE' involving the universal kriging option of Olea will be examined separately. The main form of random-to-grid interpolation available was the pointwise, distance-weighted averaging routine 'GRID' and its associated search routines 'NEAR', 'QUADRANT', 'OCTANT' and 'VRADIUS'. Where possible, all routines were used in their most basic default form. Grid estimation therefore involved a distance-weighted average of nearby sample points and not a weighted average of surfaces fitted to the local data. Within this interpolation, SURFACE II GRAPHICS allows the user to vary the search method and the weighting function used in the subsequent estimation process.

The effect of varying these parameters was examined in detail. Each search procedure (NEAR, OCTANT, QUADRANT and VRADIUS) was used with each data set and the weighting factor, 'WO'. Additionally, each weighting factor (WO,1,2,3 and 4) was utilised in conjunction with the NEAR search procedure and the

scattered data sets. Where grid nodes could not be evaluated due to insufficient data points being found by the search routine in operation, a value of -99.999 was established. In general, such nodes were not in the border area around the model and therefore were not a direct result of 'EDGE EFFECTS' (poor interpolation on the edges caused by the majority of the search area being outside the area of interest and thus of the available data points - resulting in an insufficient number of data points being found). It is important to note where this has occurred, since it has meant that extreme values, which may have occurred in the ORIGINAL-MIN and -MAX in other interpolation methods' statistics, have been omitted from all SURFACE II GRAPHICS ORIGINAL statistics. To that extent, these statistics are biassed.

6.2.1.1 DEFAULT

The DEFAULT search routine within GRID is NEAR which, for each grid node, selects (by default) the nearest 8 points within the maximum default radius limits. The 'searched' points are weighted and averaged and must therefore always produce an interpolated value within the range of the data points and thus a smoother surface than the original data surface. The results of using DEFAULT are found in Table 6.3a/b.

Examining the consistency parameters (%BAD, ORIGINAL MIN and MAX), it is apparent that the DEFAULT search method has very variable consistency. The %BAD statistics vary quite abruptly from 0.0 to 26.96 (one quarter of the values not

DS	%BAD	ORIGINAL		TRIMMED		TRIMMED ERROR				TRIMMED ABSOLUTE ERROR			CORRELATIONS			
		MIN	MAX	MIN	MAX	MEAN	MEDIAN	SKEWNESS	KURTOSIS	MEAN	MEDIAN	STANDARD	SURFACE	NEAR	NONN	GRADIENT
IB	1.21	(-34.94	36.39)	-16.41	16.33	1.10	0.79	0.06	3.39	4.46	3.22	5.93	0.991	0.313	-0.230	0.242
IC	0.17	(-19.02	27.63)	-10.60	10.55	0.06	0.06	0.08	2.73	3.32	2.73	4.19	0.995	0.123	0.021	0.114
IG	0.00	-20.33	21.39	-8.97	8.96	0.03	-0.03	-0.05	3.14	2.42	1.85	3.26	0.997	0.525	-0.452	0.255
IR	5.82	(-71.97	33.80)	-27.33	26.14	-5.39	-4.61	-0.10	3.89	7.68	6.05	9.97	0.984	0.472	-0.357	0.314
IS	0.00	-20.89	38.74	-13.36	13.35	0.62	0.82	-0.22	3.42	3.65	2.81	4.80	0.994	0.224	-0.060	0.312
LB	10.40	(-8.05	7.65)	-4.47	4.47	0.06	0.05	0.00	2.79	1.40	1.14	1.77	0.997	0.227	-0.161	0.141
LC	3.84	(-8.67	7.31)	-5.56	5.72	0.33	0.03	0.52	3.10	1.47	1.17	1.80	0.996	0.357	-0.304	0.214
LG	0.00	-7.88	4.72	-2.59	2.63	0.00	-0.12	0.03	3.17	0.73	0.58	0.96	0.999	0.452	-0.079	0.308
LR	26.96	(-14.81	12.97)	-6.01	5.84	-0.96	-0.77	-0.24	4.43	1.39	0.99	1.88	0.995	0.389	-0.302	0.382
LS	0.14	(-12.18	8.97)	-4.15	4.23	0.00	-0.08	0.18	3.36	1.14	0.89	1.48	0.998	0.068	0.069	0.345
FB	6.43	(-7.65	10.17)	-5.54	5.67	0.54	0.48	0.06	4.26	1.23	0.88	1.68	0.953	0.272	-0.211	0.173
FC	0.00	-7.52	6.47	-3.48	3.50	0.63	0.67	-0.42	3.60	1.14	0.94	1.43	0.971	0.141	0.070	0.077
FG	0.00	-10.11	4.97	-2.60	2.56	0.21	0.21	-0.34	3.81	0.65	0.50	0.87	0.987	0.371	-0.048	0.187
FS	0.05	(-8.89	11.02)	-4.93	5.13	0.79	0.54	0.43	3.57	1.27	0.88	1.75	0.954	0.266	-0.182	0.169

Table 6.3a Primary results of SURFACE II GRAPHICS (DEFAULT)

DS	IN-NEAR			IN-NONN			IN-GRAD			OUT-NEAR			OUT-NONN			OUT-GRAD		
	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN
IB	0.95	174.83	50.36	0	14	3.82	0.12	83.55	19.07	18.83	144.69	75.17	0	9	1.77	4.38	83.00	25.41
IC	0.83	182.47	43.61	0	10	3.32	0.12	83.55	18.92	10.31	157.25	57.58	0	9	3.32	1.00	83.00	27.66
IG	0.00	42.42	32.19	4	6	4.21	0.17	83.00	18.78	30.00	42.42	37.03	4	5	4.04	0.12	83.55	30.64
IR	0.91	196.77	65.82	0	20	3.59	0.17	83.55	18.61	30.00	306.41	176.08	0	5	0.22	0.12	78.75	27.53
IS	0.31	148.00	45.63	0	12	3.36	0.12	83.55	18.69	21.83	150.02	68.69	0	6	2.39	5.32	78.75	32.46
LB	0.31	105.99	32.02	0	23	3.60	0.00	41.50	10.85	9.48	165.00	84.28	0	8	0.60	0.00	29.74	7.86
LC	0.00	106.07	25.97	0	11	3.40	0.00	41.50	10.71	9.02	133.99	75.24	0	6	0.53	0.35	26.72	4.47
LG	0.00	42.42	19.58	1	6	3.35	0.00	41.50	10.00	15.00	42.42	24.11	1	5	2.97	2.69	39.77	18.05
LR	0.31	103.08	35.94	0	19	4.43	0.00	41.50	8.39	3.26	226.99	110.36	0	12	0.36	0.50	35.99	15.22
LS	0.21	105.08	27.19	0	11	3.06	0.00	41.50	10.02	10.34	112.92	36.86	0	9	2.60	0.74	33.14	17.58
FB	0.21	82.46	24.91	0	18	5.02	0.00	84.47	8.72	10.88	129.47	75.86	0	12	0.46	0.13	55.25	7.31
FC	0.40	70.00	21.36	0	17	4.55	0.00	84.47	8.42	5.47	48.39	22.38	1	15	6.47	0.13	55.25	12.17
FG	0.00	41.22	16.27	2	7	4.76	0.00	84.47	7.96	10.00	41.22	18.80	2	7	4.60	1.07	74.47	20.97
FS	0.00	80.61	24.08	0	17	4.68	0.00	84.47	8.50	8.46	84.33	34.86	0	12	3.96	0.13	55.25	10.68

Table 6.3b Secondary results of SURFACE II GRAPHICS (DEFAULT)

estimated). This is clearly undesirable although it is explained by the highly significant rank correlation (0.86) between %BAD and NNS (see Table 3.2). This is supported by considering the IN-NEAR and OUT-NEAR statistics. Where high %BAD statistics occur - most noticeably with the RIVER and BREAKLINE data sets - there is a dramatic increase in the NEAR MEAN (OUT minus IN). Indeed, over the full range of data sets, this amounts to a very highly significant rank correlation of 0.92. Alternatively, only a significant rank correlation (-0.83) exists between %BAD and the NONN MEAN (OUT minus IN), suggesting that node/data point nearness is the main contributor to poor interpolation consistency.

The ORIGINAL MIN and MAX statistics also provide a measure of bias of the interpolation. In general, the magnitudes of the MIN and MAX values are similar, suggesting the well-balanced nature of the interpolation. Notable exceptions are INCH-RIVER and -SCATTER which must indicate special features of the data sets.

A similar strong correlation occurs between NNS and the accuracy desciptors. STANDARD ERROR and ABS MEAN have rank correlations of 0.75, although only SURFACE CORRELATION produces a significant rank correlation of 0.82 with NNS. Near values tend to be similar and distant values differ, therefore if the nearest data point to a grid node is quite distant, the chances are that the interpolated value could be strongly in error. This manifests itself in a weak, yet relatively consistent ERROR/NEAR correlation - mean of 0.300. It is notable that a low negative correlation exists between error

and the number of near neighbours (ERROR/NONN). It is clearly important to have a few neighbours to provide a reasonable local sample, although a large number is not as important as is an even distribution in the locality of the grid node.

Noticeably, the next strongest accuracy-related rank correlations occur with the grid-based areal autocorrelation function (see Table 4.3). Moderate rank correlation of 0.67 occurs with SURFACE CORRELATION and ABS MEAN and 0.73 occurs with STANDARD ERROR. The GRID interpolation is a data averaging process. Where the surface is smooth and flat, the data points to be averaged will have similar characteristics, providing a similar interpolated value. However, when the surface is rough, or even when there is a constant terrain slope, and the data points found by the search routine do not represent a balanced height distribution in the environment of the interpolated point, the value will be in error to an extent related to the general trend of the local surface and the point distribution.

Examining the STANDARD ERROR in more depth, it is noticeable that this form of interpolation provides generally consistent results when the point distribution is reasonably even. Serious problems do occur where the surface is particularly variable and the data points are poorly distributed - INCH-RIVER (NNS of 0.804) and INCH-BREAKLINE (NNS of 1.093). Furthermore, where a data point is located on a grid node, it is adopted as the interpolated value; clearly if there are many such instances, the GRID grids will have a particularly low STANDARD ERROR. Excluding these special

cases, generally low within-model variations are produced by this method.

Another interesting weak, yet consistent correlation exists exists between error and gradient (ERROR/GRAD). This is a direct result of a skewed data point height distribution within an individual search area having a maximum effect when the gradient is at a maximum. For example, on a constantly rising slope, a simple weighted mean will systematically interpolate low if the majority of points are at the bottom of a slope and few points are at the top of a slope. The magnitude of the error will vary directly with the steepness of the slope. Therefore, excluding RIVER data (which in terms of surface representation have an inherently positively skewed height distribution), the surface-specific data sets (BREAKLINE and CONTOUR) produce weaker correlation coefficients than do the surface-random data sets (GRID and SCATTER). Use of FORV data results in much weaker correlation coefficients due to their less-normalised distribution of slope across the surface - positively skewed with a high kurtosis (see Table 4.10b).

The MEAN and MEDIAN ERROR, TRIMMED MIN and MAX and SKEWNESS provide descriptors of systematic error and bias in the interpolation method. In general, systematic error, as portrayed by the MEAN ERROR, is small in these data sets. A low average in one area will be compensated by a high average in another, thus maximum and minimum errors are approximately equal and, excluding RIVER data, low mean errors exist (less than +/-1.10m). It is not surprising that RIVER data should

have high negative means. Rivers occupy the lowest elevations and, since the interpolation method averages the data values, RIVER data must always produce systematically low results. More interestingly, CONTOUR data generate a small positive systematic error, while GRID data consistently produce low systematic error. SCATTER data produce a variable positive systematic error and BREAKLINE data produce generally variable results.

Surprisingly, the MEAN and MEDIAN ERRORS have only a weak rank correlation (0.50) with the mean altitudes (see Table 4.10a) as derived from the surface characteristic (pointwise polynomial) program GEOD. While this may be a facet of the indirect method of generating the mean altitudes, it may also be due in part to the effect of the altitudinal range being partially obscured by the data distribution. For example, there are large voids between RIVER data points which, if they are systematically interpolated low, will generate a large negative error. CONTOUR and GRID data sets are, by contrast, well distributed in the plane, therefore localised negative and positive error areas will be small, randomly distributed and tend to cancel each other out. BREAKLINE and SCATTER data sets have a more heterogeneous distribution of data void sizes which results in a more random distribution of systematic error.

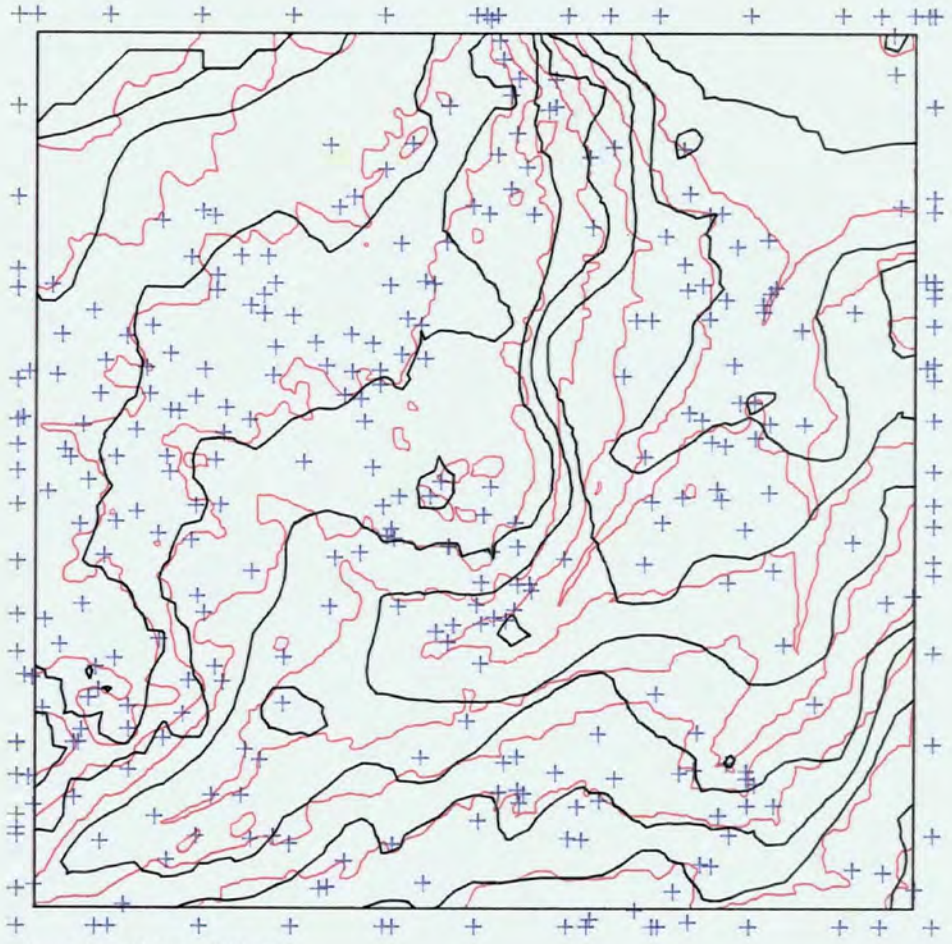
In general, there is little variation between the use of the MEAN and MEDIAN ERROR statistics. Indeed, a highly significant rank correlation of 0.86 exists between MEAN and MEDIAN ERROR - a perfect 1.0 correlation existing between ABS

MEAN and ABS MEDIAN. Major differences are naturally related to large SKEWNESS (e.g. in the FORV-SCATTER and LIAN-CONTOUR data sets) but, since this is generally modest (as discussed previously), negative and positive errors tend to correspond.

The precision statistics (TRIMMED MIN and MAX and KURTOSIS) provide another useful insight into the interpolation method. KURTOSIS is remarkably consistent throughout all data sets, being generally slightly more peaked than normal. This method is therefore a moderately safe interpolator, generating moderately high precision results. This is supported by the TRIMMED MIN and MAX which are of generally low magnitude (for LIAN and FORV) and well-balanced. INCH noticeably produces greater within-model variance, underlining the link between varied NNSs occurring with a more varied relief. For example, the variation in relative magnitude for INCH-GRID and INCH-RIVER is approximately 3.5 whereas for LIAN it is only 2.5.

In summary, therefore, while examining the DEFAULT sections of Figures 6.1, 6.2 and 6.3, little systematic error exists across any of the models. Areas where reference isarithms are outside computer isarithms are compensated by areas where the reference isarithms are inside. The strong correlation between error and nearest neighbour distance (ERROR/NEAR) is best illustrated in Figure 6.1a by INCH-BREAKLINE. Where a dense point distribution exists, the fit is good, although in sparse areas, notably top-left and top-right-of-centre, the fit is very much poorer. This is obvious in Figure 6.2a where the even distribution of INCH-GRID produces superior results in

Figure 6.1



I-B DEFAULT

I-B OCTANT

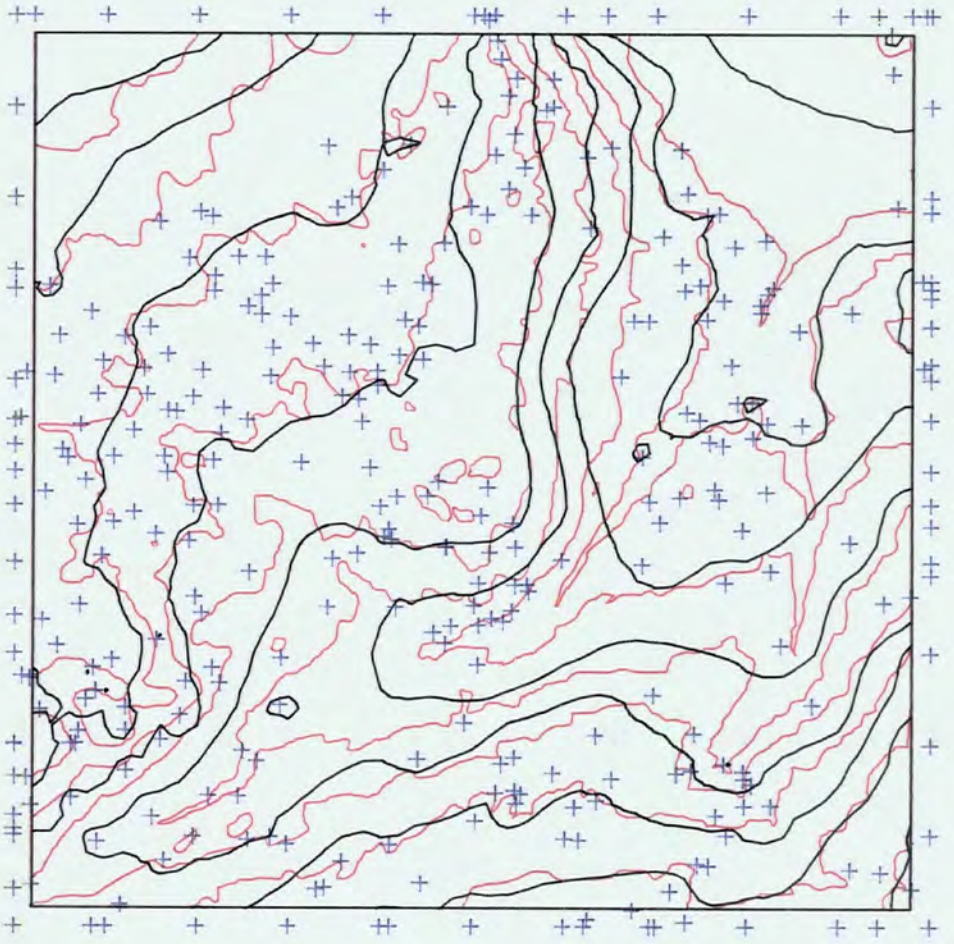
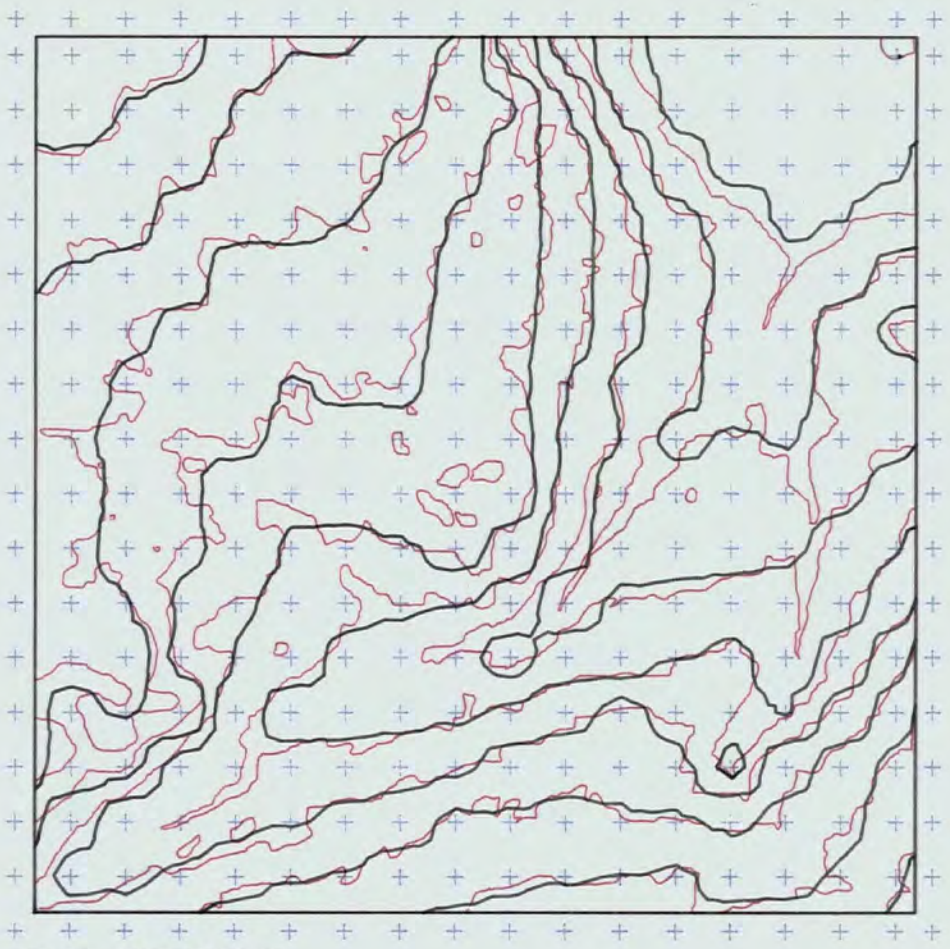


Figure 6.2



I-G DEFAULT

I-G VRADIUS

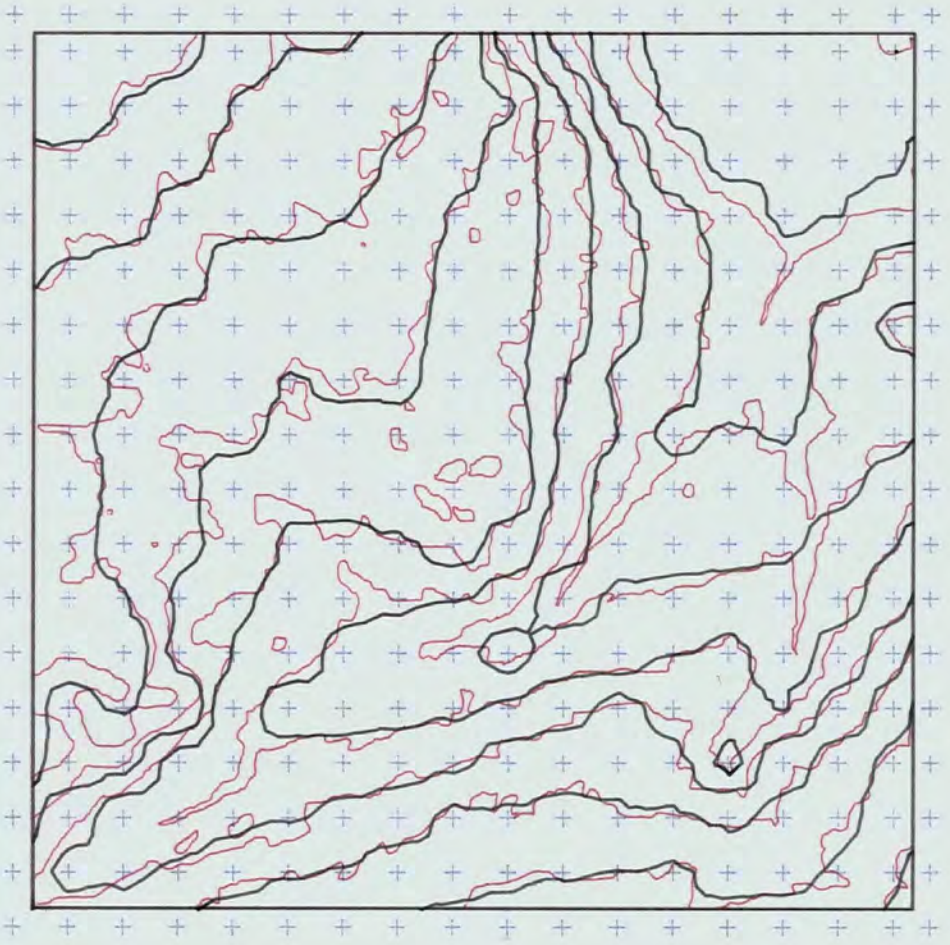
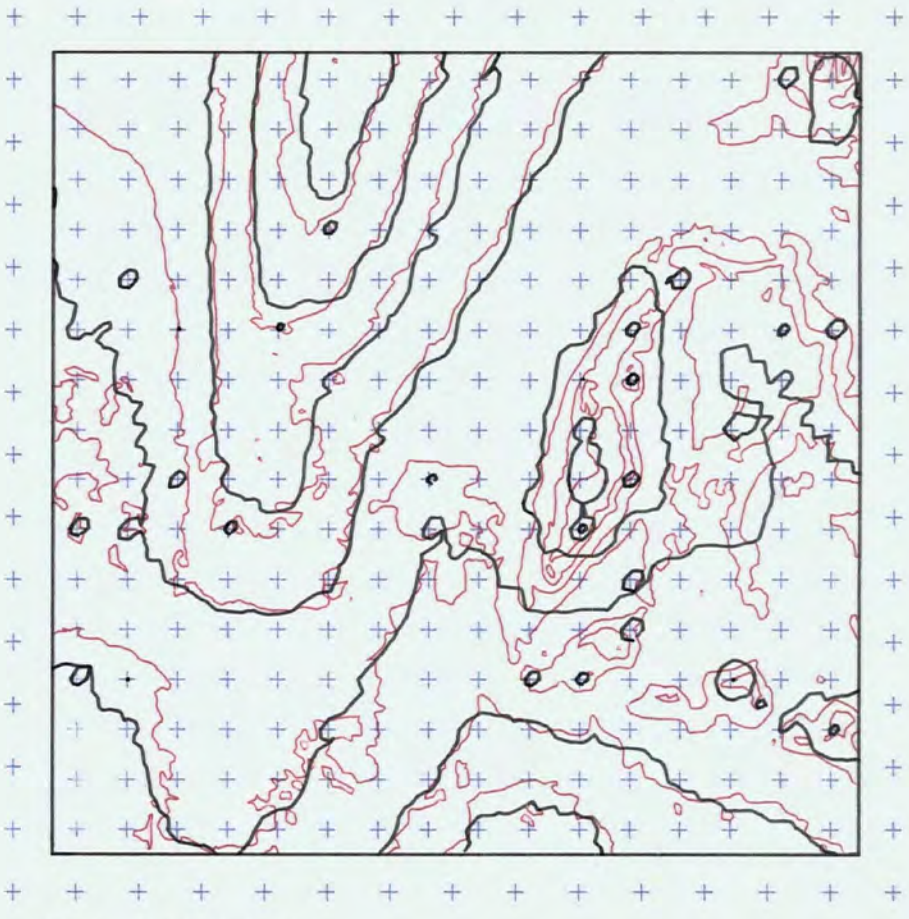
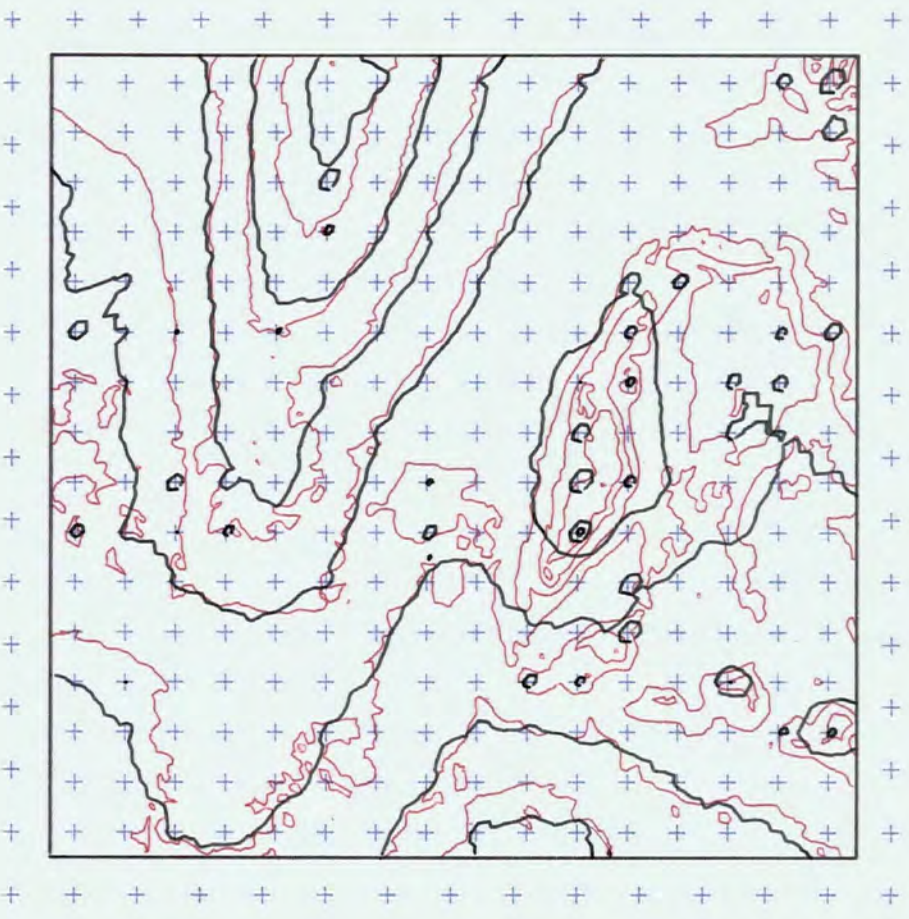


Figure 6.3



F-G DEFAULT

F-G VRADIUS



those areas especially, although a similarly regular distribution in FORV-GRID (Figure 6.3a) produces significantly poorer results while still preserving the general shape of the terrain. The latter distortion is an illustration of the correlations between surface autocorrelation and error.

An additional feature of using FORV data within SURFACE II GRAPHICS is illustrated in Figure 6.3 and subsequently in Figures 6.5 and 6.6. Because SURFACE II GRAPHICS adopts values of data points for grid nodes where the two coincide, areas of consistently low or high interpolated values, often in areas of sparse data, are interspersed with small circular isarithms round adopted points outside the range of the surrounding interpolated environment. There is a liberal distribution of such 'ADOPTED ISARITHMS' in Figures 6.3 and 6.5 which, if the surrounding interpolated and reference isarithms are examined, will be seen to be in such situations.

6.2.1.2 QUADRANT

The search process may be 'directionalised' by using the QUADRANT search routine. In order to avoid corner problems, where the searchable area might be restricted to the extent that no corner nodes in the grid would be evaluated, this was defined such that a minimum of one quadrant had to have a successful search for the node to be evaluated. A maximum of four points were sought within each quadrant; within each quadrant, the closest point had to be within a fixed radius and all four points within a maximum radius. The fixed radii were 600m and 1000m (INCH), 375m and 650m (LIAN) and 180m and

500m (FORV). These radii were chosen to reflect the relative scales and grid resolutions of the three models.

The general effect of utilising the QUADRANT (see Table 6.4) as opposed to the NEAR (DEFAULT) search, is to maintain or slightly improve the majority of error statistics. However, on closer examination, this is shown to be restricted mainly to the INCH and LIAN data sets, probably due to the low autocorrelation within FORV.

The most important improvements are in the consistency statistics. Virtually all of the nodes poorly interpolated (%BAD) have been removed. Without exception, all unsuccessful searches using QUADRANT are a subset of, and occur at, the unsuccessful nodes in the DEFAULT searches. There is still, however, a clear relationship between %BAD and the NNSs since only one data set in each model has a non-zero %BAD value, and that data set also has the lowest NNS in that model.

An important aspect of the ORIGINAL MIN and MAX statistics is the lack of failed searches (signified by bracketing the values), suggesting a more easily satisfied minimum search requirement. If more searches were marginally satisfied, it would be expected that the subsequent averaging process would use more distant points; hence the interpolated result would be in considerable error, increasing the ORIGINAL MIN and MAX. In general, this has not happened since the magnitude of the MINs and MAXs have declined (with the exception of a few trivial cases), while preserving similar relationships between the magnitude of the MIN and MAX values - mean values for the NEAR search of -18.07 and 16.59 have declined to -16.72 and

DS	%BAD	ORIGINAL		TRIMMED		TRIMMED ERROR				TRIMMED ABSOLUTE ERROR			CORRELATIONS			
		MIN	MAX	MIN	MAX	MEAN	MEDIAN	SKEWNESS	KURTOSIS	MEAN	MEDIAN	STANDARD	SURFACE	NEAR	NONN	GRADIENT
IB	0.00	-29.75	35.19	-14.36	14.40	0.79	0.59	0.05	3.04	4.26	3.36	5.52	0.993	0.323	-0.239	0.150
IC	0.00	-18.99	25.39	-11.00	11.05	0.22	0.22	0.15	2.88	3.31	2.71	4.22	0.995	0.144	-0.004	0.069
IG	0.00	-18.47	21.58	-8.90	8.88	0.04	-0.09	-0.04	3.31	2.31	1.72	3.14	0.997	0.525	-0.447	0.208
IR	0.87	-61.63	19.92	-24.00	18.89	-5.98	-5.14	-0.02	3.56	7.27	5.88	9.27	0.988	0.438	-0.320	0.289
IS	0.00	-21.35	35.41	-13.19	13.05	0.49	0.69	-0.19	3.39	3.43	2.64	4.47	0.995	0.193	-0.014	0.257
LB	0.00	-6.42	5.94	-3.89	3.89	0.30	0.41	-0.25	2.72	1.31	1.10	1.64	0.997	0.227	-0.141	0.092
LC	0.00	-7.26	7.09	-4.23	4.23	0.39	0.16	0.33	2.57	1.35	1.10	1.72	0.997	0.293	-0.235	0.192
LG	0.00	-7.84	4.19	-2.39	2.42	0.01	-0.09	-0.04	3.06	0.69	0.55	0.90	0.999	0.463	-0.083	0.271
LR	0.04	(-13.81	17.19)	-9.77	9.81	-0.82	-0.76	0.52	5.64	2.02	1.26	2.90	0.992	0.536	-0.422	0.498
LS	0.00	-11.48	8.77	-3.96	3.97	0.03	-0.03	0.12	3.14	1.10	0.88	1.43	0.998	0.072	0.060	0.330
FB	0.33	(-9.27	8.51)	-5.17	5.21	0.71	0.69	-0.16	3.57	1.48	1.14	1.93	0.936	0.380	-0.344	0.058
FC	0.00	-8.14	6.22	-3.64	3.64	0.73	0.85	-0.56	3.44	1.26	1.13	1.54	0.963	0.188	-0.030	-0.038
FG	0.00	-10.72	4.30	-2.93	2.92	0.24	0.31	-0.54	3.44	0.84	0.68	1.09	0.978	0.373	-0.005	0.120
FS	0.00	-9.02	9.59	-4.52	4.93	0.75	0.50	0.43	3.60	1.21	0.82	1.65	0.959	0.335	-0.198	0.207

Table 6.4a Primary results of Surface II Graphics (QUADRANT)

DS	IN-NEAR			IN-NONN			IN-GRAD			OUT-NEAR			OUT-NONN			OUT-GRAD		
	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN
IB	0.95	174.83	50.58	0	14	3.80	0.12	83.55	18.88	18.63	144.60	70.88	0	9	2.31	1.22	83.00	28.80
IC	0.83	182.47	43.32	0	10	3.34	0.12	83.55	19.11	17.74	157.25	63.80	0	8	3.09	1.00	83.00	24.30
IG	0.00	42.42	32.17	4	6	4.21	0.17	83.00	18.88	30.00	42.42	37.25	4	5	4.04	0.12	83.55	28.88
IR	0.91	306.41	70.53	0	20	3.46	0.12	83.55	18.66	66.78	297.06	168.25	0	5	0.18	3.81	78.75	32.88
IS	0.31	150.02	45.63	0	12	3.36	0.12	83.55	18.77	21.83	131.77	68.74	0	7	2.40	2.90	77.96	30.86
LB	0.31	165.00	39.41	0	23	3.21	0.00	41.50	10.19	7.34	131.83	51.16	0	20	1.72	0.76	29.74	14.26
LC	0.00	133.99	27.13	0	11	3.36	0.00	41.50	10.65	9.02	117.52	53.19	0	8	1.25	0.50	27.72	5.47
LG	0.00	42.42	19.55	1	6	3.35	0.00	40.00	10.05	15.00	42.42	24.77	1	4	2.96	2.69	41.50	17.22
LR	0.31	226.99	53.63	0	19	3.42	0.00	41.50	10.06	32.64	202.50	135.42	0	1	0.02	6.17	30.19	16.83
LS	0.21	112.92	27.41	0	11	3.05	0.00	41.50	10.01	4.86	63.55	33.00	0	9	2.89	0.74	33.14	17.86
FB	0.21	129.47	27.22	0	18	4.84	0.00	84.47	8.68	9.29	129.38	67.14	0	12	0.96	0.13	56.39	7.26
FC	0.40	70.00	21.41	0	17	4.56	0.00	84.47	8.40	2.47	41.97	21.25	1	15	6.34	0.56	56.39	12.42
FG	0.00	41.22	16.27	2	7	4.76	0.00	84.47	7.97	10.00	41.22	18.78	2	7	4.65	0.91	74.47	20.72
FS	0.00	84.33	24.03	0	17	4.68	0.00	84.47	8.46	7.42	76.17	35.88	0	12	3.96	0.56	55.25	11.51

Table 6.4b Secondary results of SURFACE II GRPAHICS (QUADRANT)

14.95 for the QUADRANT search.

The accuracy statistics similarly display modest improvements for all INCH and LIAN data sets (with the exclusion of LIAN-RIVER). For these nine data sets, the mean STANDARD ERROR has improved from 3.09 to 2.95, the mean SURFACE CORRELATION from 0.9945 to 0.9954 and the mean ABS MEAN from 2.92 to 2.79. This suggests that, for these data sets, areal autocorrelation is such that, even though more distant points are perhaps being selected, they are at least as representative of the local surface as those found by the NEAR search.

As in other analyses, those using the FORV data sets differ: they show a modest decrease in accuracy with use of the OCTANT option. The FORV mean STANDARD ERROR has decayed from 1.56 to 1.68, the FORV mean SURFACE CORRELATION has decayed from 0.966 to 0.959 and the FORV mean ABS MEAN has decayed from 1.08 to 1.20. Since all the FORV data sets have moderate to high NNSs, this cannot be directly associated with data distribution. It is most likely due to the rapid decay in the FORV autocorrelation functions, associated with the fact that the more directional QUADRANT search finds more distant points which therefore provide a more distorted sample for the averaging process.

Results based on the LIAN-RIVER data set also show a considerable diminu tion of accuracy. This is more associated with its low NNS (0.768) which results in the directionalised search finding extremely distant points for many more nodes. This is best demonstrated by the ERROR/NEAR and ERROR/NONN

correlations which are generally of a lower strength for the QUADRANT search. However, for LIAN-RIVER they show a marked increase (NEAR - 0.389 to 0.536, NONN - -0.302 to -0.422), thus stressing the importance of the point distribution.

The associations between accuracy and NNS and autocorrelation are underlined by considering the relevant rank correlations. SURFACE CORRELATION has a very highly significant (0.93) correlation, STANDARD ERROR has a highly significant (0.89) correlation and ABS MEAN has a significant correlation with NNS. It is worth noting that the next highest rank correlations are with the autocorrelation functions, although these only account for about 50% of the variance.

The biassing, precision and systematic error descriptors behave in a similar fashion to the DEFAULT statistics, although with a generally reduced magnitude. In total, the TRIMMED MIN and MAX statistics produce a 0.98 rank correlation and the MEAN error has a 0.99 correlation between DEFAULT and QUADRANT. The average magnitude of the TRIMMED MIN and MAX has improved from 8.26 to 7.83 and the average MEAN error has changed from -0.14 to -0.15 by applying the QUADRANT search. Hidden within these global statistics are the large increases in the TRIMMED MIN and MAX values for LIAN-RIVER, and the increased negative systematic error for the RIVER data. Overall, then, QUADRANT searches reduce the range of error without seriously affecting the systematic error. However, this does not result in a peaking of the error distribution around the mean, as the mean KURTOSIS has dropped from 3.48 to 3.39 despite a large increase for LIAN-RIVER . Taking the reduced

error range into account, this suggests that the error distribution has been spread away from the mean as much as it has been focussed onto the mean. This is understandable, since in situations - especially with lineated data - where perhaps eight data points lie very close to a grid node within one quadrant, QUADRANT will select only four of these points along with other, more distant (and possibly more different) points in other quadrants. Alternatively, where eight clustered, but distant, unrepresentative data points were selected by the NEAR algorithm, a more spatially balanced selection of distant points will be found using the QUADRANT search. The first scenario increases the error in low error situations (in general where MEAN NEAR errors were less than +/-0.5), whereas the second scenario decreases the error in high error situations (in general where the MEAN NEAR errors were greater than +/-0.8).

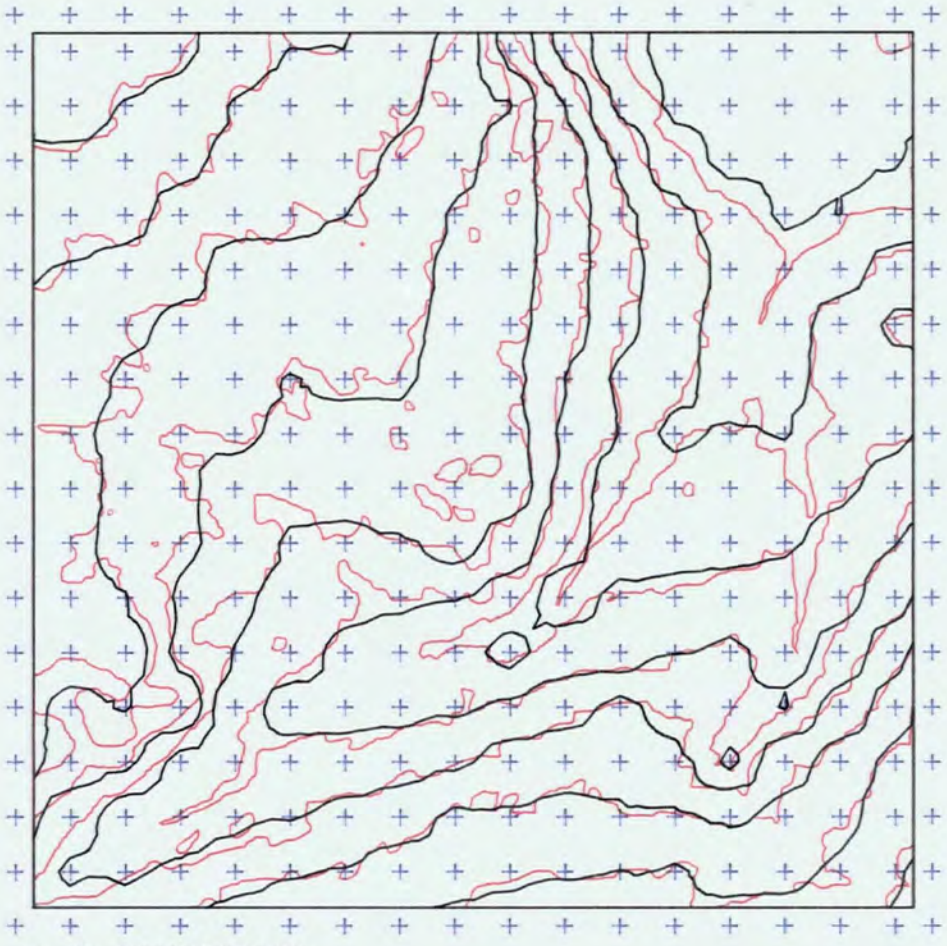
Interestingly, the magnitudes of the KURTOSIS changes are inversely related to terrain variability as defined by vector dispersion (Table 4.8). Generally, QUADRANT searching has increased the distance between the data points used to interpolate a grid node and the grid node. Intrinsically, this further spatial averaging produces 'safe' but far-from-optimal interpolations. Other aspects of the error distribution are unchanged. While specific values fluctuate, SKEWNESS remains broadly similar. Additionally, the relative relationship between MEAN and MEDIAN error is unchanged, as is the relationship between ABS MEAN and MEAN.

Error-related correlations have undergone changes which

reinforce the previous discussion. Where data sets are sufficiently autocorrelated, the ERROR/NEAR and ERROR/NONN correlations have moderately declined. The QUADRANT method is using more distant points, although due to the sufficiently high autocorrelation, these points are similar enough, and may be balanced by other distant points. However, where the autocorrelation decay is more rapid the error can increase more rapidly with distance. This has resulted in a moderate increase in the magnitude of the total mean ERROR/NEAR (0.300 to 0.321) and ERROR/NONN (-0.159 to -0.173), although it is additionally caused by the extremely poor NNS of LIAN-RIVER. Alternatively, the ERROR/GRAD correlations have consistently declined. Since more distant data points are found in the QUADRANT search than in the DEFAULT search, the effect of the immediate local gradient is not as important as that of the surrounding area.

In general, while considering Figures 6.4a, 6.5a and 6.6b, little major improvement may be seen from the resultant isarithms, with any changes between the DEFAULT and QUADRANT methods being balanced by an equal and opposite change elsewhere. Improvements occur in the flatter areas, for example the left border of FORV (Figure 6.5a and 6.6b), while increases in error occur in the steeper areas (right-of-centre in FORV). Because of the high autocorrelation present in the INCH data, few changes and errors occur with the resulting isarithms (Figure 6.4), while use of FORV data, with its associated low autocorrelation, produces larger changes and errors (Figures 6.4 and 6.5).

Figure 6.4



I-G QUADRANT

I-G OCTANT

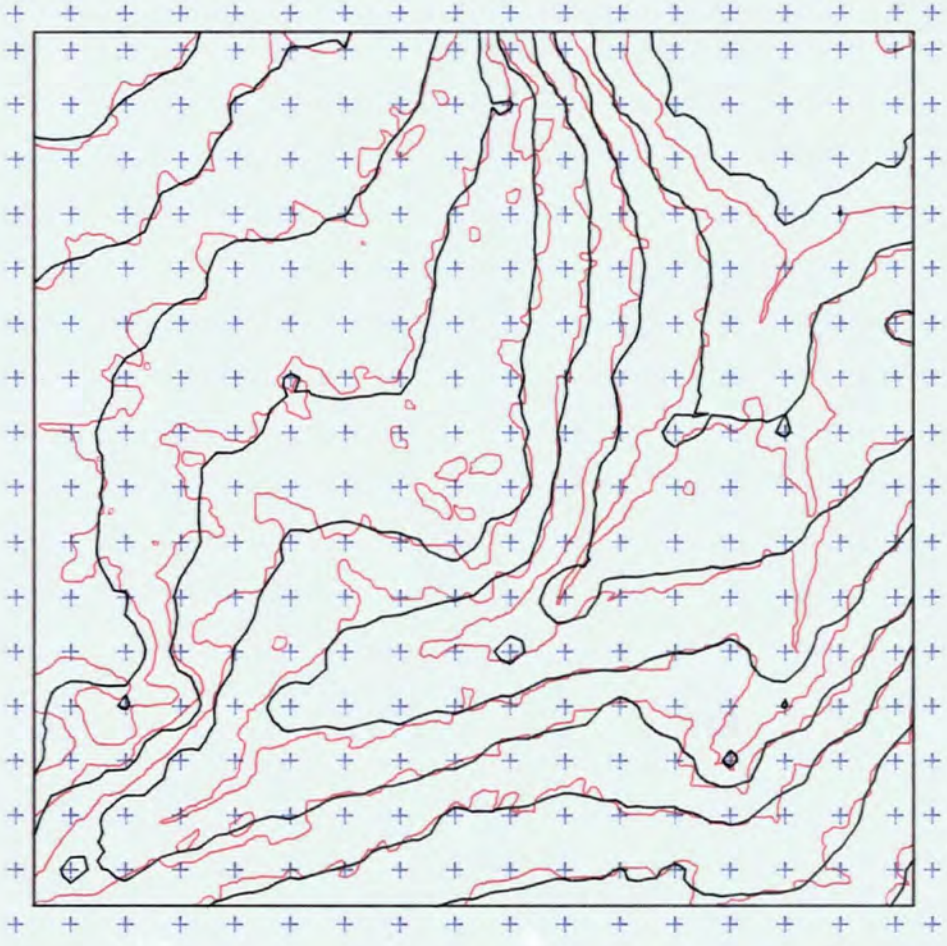
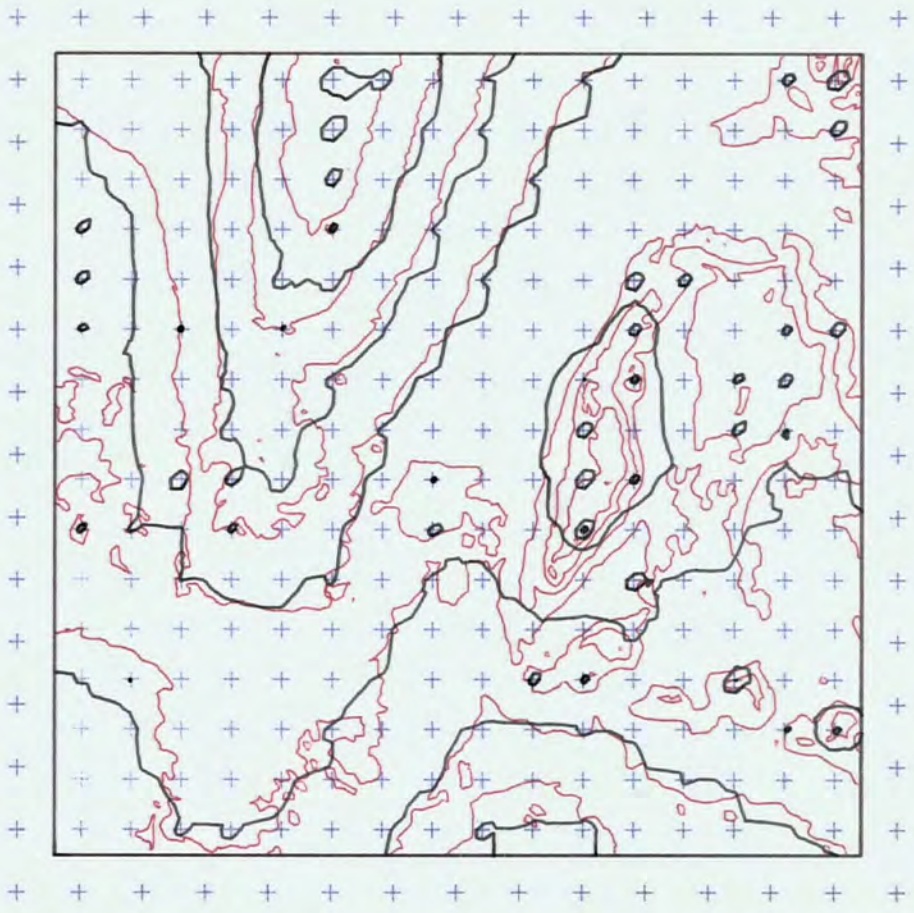


Figure 6.5



F-G QUADRANT

F-G OCTANT

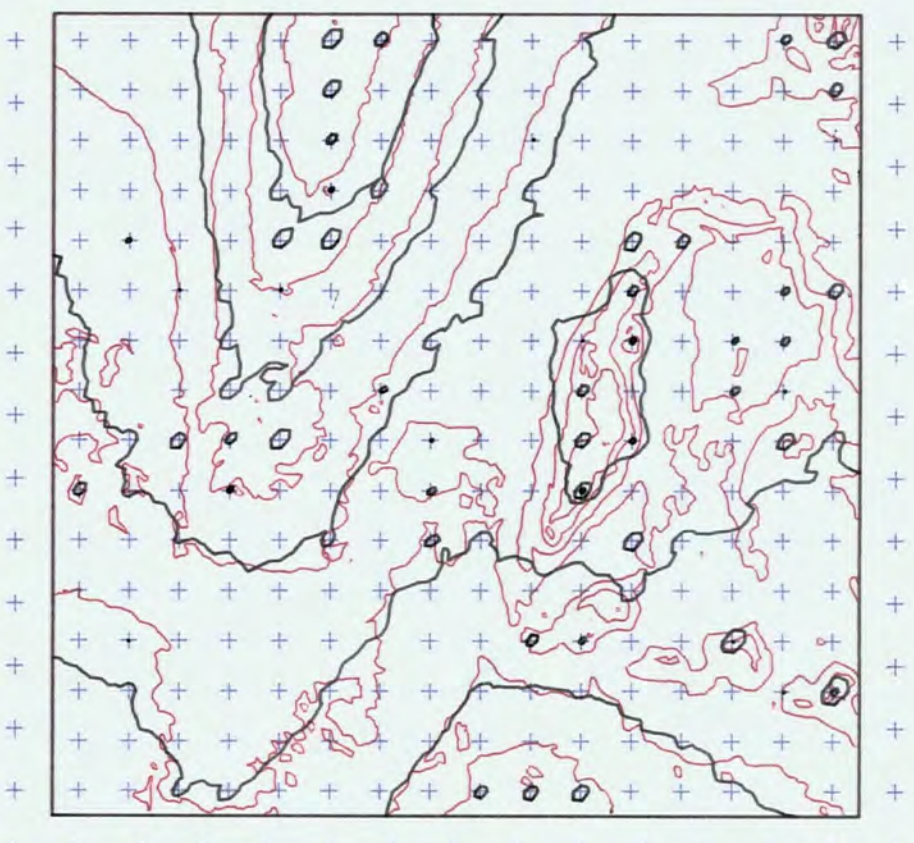
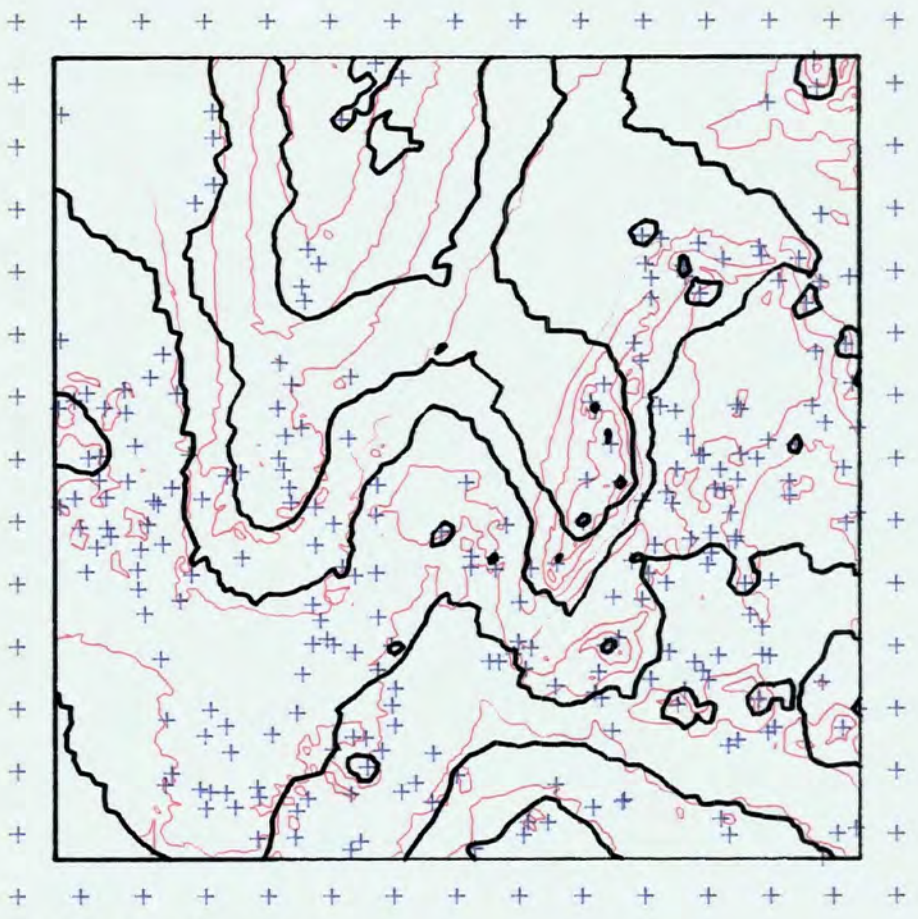
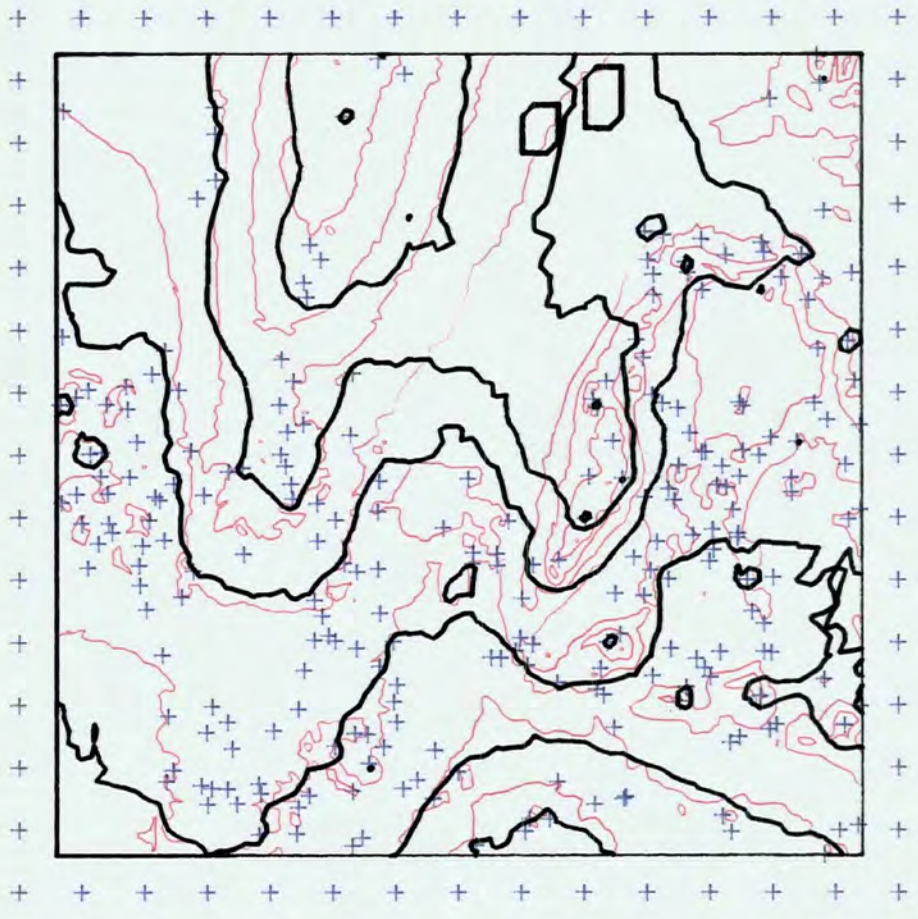


Figure 6.6



F-B VRADIUS

F-B QUADRANT



6.2.1.3 OCTANT

The directional QUADRANT search may be further refined using an OCTANT search. This search was defined such that a minimum of two octants had to have a successful search for the node to be evaluated. A successful search was defined by a maximum of four points per octant and, within this, the closest and most distant points had to be within a fixed range, 600m and 1000m (INCH), 375m and 650m (LIAN) and 300m and 590m (FORV). These ranges must be greater than for QUADRANT as each search segment is halved and will thus contain half the number of points. To find the same number of points within each segment, the ranges must be increased accordingly.

The results of utilising the OCTANT search (see Table 6.5) generally lie between the DEFAULT and QUADRANT results but are closer to the latter. Like QUADRANT, there is a distinct improvement in the number of nodes poorly interpolated (%BAD) although, with the exclusion of RIVER data, the undetermined nodes are not related to those found in the QUADRANT or DEFAULT search. This is due to a combination of increased search radii and smaller search segments. Assume that, around one grid node, three quadrants contain no points and, in the other quadrant, three points occur in one octant and one point occurs in the other octant. While the QUADRANT search will be successful, the OCTANT search around the grid node will fail. Assume also that around another grid node, four points occur in each of two contiguous octants which are in different

DS	%BAD	ORIGINAL		TRIMMED		TRIMMED ERROR				TRIMMED ABSOLUTE ERROR			CORRELATIONS			
		MIN	MAX	MIN	MAX	MEAN	MEDIAN	SKEWNESS	KURTOSIS	MEAN	MEDIAN	STANDARD	SURFACE	NEAR	NONN	GRADIENT
IB	0.00	-26.27	38.91	-14.73	14.71	0.69	0.44	0.07	2.90	4.54	3.48	5.85	0.992	0.354	-0.271	0.107
IC	0.00	-20.41	28.85	-12.04	12.09	0.23	0.20	0.22	3.05	3.47	2.73	4.46	0.995	0.163	-0.010	0.063
IG	0.00	-19.16	22.78	-9.89	9.88	0.11	-0.09	0.01	3.52	2.40	1.77	3.31	0.997	0.517	-0.445	0.190
IR	1.04	-61.71	19.36	-21.80	19.07	-6.22	-5.63	-0.02	3.25	7.15	6.13	8.96	0.990	0.451	-0.319	0.293
IS	0.00	-20.75	36.58	-13.43	13.21	0.39	0.54	-0.13	3.31	3.50	2.79	4.52	0.995	0.198	-0.001	0.204
LB	0.17	(-6.57	6.18)	-3.82	3.84	0.44	0.58	-0.35	2.68	1.34	1.14	1.61	0.997	0.270	-0.159	0.021
LC	0.03	(-6.69	6.93)	-4.25	4.32	0.53	0.32	0.25	2.52	1.39	1.14	1.77	0.997	0.300	-0.227	-0.178
LG	0.00	-7.92	4.17	-2.42	2.40	0.03	-0.08	-0.10	3.10	0.69	0.54	0.90	0.999	0.470	-0.084	0.248
LR	0.93	(-10.35	14.69)	-8.02	8.00	-0.89	-0.81	0.41	5.13	1.81	1.18	2.54	0.994	0.493	-0.386	0.489
LS	0.00	-7.68	8.29	-3.18	3.07	0.20	0.18	-0.07	2.68	1.05	0.88	1.30	0.998	0.091	0.050	0.240
FB	0.17	(-10.07	7.88)	-5.35	5.47	0.86	1.00	-0.44	3.27	1.77	1.52	2.18	0.912	0.403	-0.399	-0.043
FC	0.00	-9.72	5.64	-4.21	4.26	0.94	1.14	-0.64	3.26	1.60	1.47	1.89	0.937	0.201	-0.084	-0.094
FG	0.00	-11.60	3.42	-3.50	3.42	0.33	0.46	-0.72	3.51	1.05	0.88	1.34	0.965	0.384	0.022	0.064
FS	0.00	-9.39	9.67	-5.13	5.21	0.89	0.67	0.21	3.48	1.42	1.01	1.94	0.947	0.300	-0.222	0.141

Table 6.5a Primary results of SURFACE II GRAPHICS (OCTANT)

DS	IN-NEAR			IN-NONN			IN-GRAD			OUT-NEAR			OUT-NONN			OUT-GRAD		
	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN
IB	0.95	174.83	50.66	0	14	3.79	0.12	83.55	18.99	18.63	155.55	69.33	0	9	2.48	1.22	83.00	26.80
IC	0.83	182.47	43.00	0	10	3.35	0.12	83.55	19.35	19.97	157.25	69.94	0	8	2.79	0.17	81.55	20.03
IG	0.00	42.42	32.16	4	6	4.21	0.17	83.00	18.97	30.00	42.42	37.46	4	5	4.04	0.12	83.55	27.08
IR	0.91	306.41	70.60	0	20	3.46	0.12	83.55	18.77	63.36	305.62	167.08	0	3	0.17	3.81	78.75	31.05
IS	0.31	150.02	45.58	0	12	3.38	0.12	83.55	18.80	26.50	131.77	69.67	0	7	2.30	2.90	77.96	30.17
LB	0.31	165.00	39.33	0	23	3.19	0.00	41.50	10.29	7.34	131.83	51.92	0	20	2.00	0.34	29.74	12.57
LC	0.00	133.99	27.13	0	11	3.35	0.00	41.50	10.64	9.02	124.67	53.36	0	8	1.35	0.50	27.72	5.68
LG	0.00	42.42	19.52	1	6	3.35	0.00	40.00	10.06	15.00	42.42	25.22	1	4	2.92	2.68	41.50	17.00
LR	0.31	226.99	54.22	0	19	3.43	0.00	41.50	9.97	18.88	202.50	113.74	0	6	0.18	5.36	29.88	17.35
LS	0.21	112.92	27.47	0	11	3.05	0.00	41.50	10.11	4.86	63.55	31.82	0	9	2.92	0.74	33.14	15.84
FB	0.21	129.47	28.13	0	18	4.81	0.00	84.47	8.67	8.84	119.25	51.63	0	12	1.35	0.13	56.39	7.47
FC	0.40	70.00	21.38	0	17	4.60	0.00	84.47	8.48	2.47	41.97	21.72	1	16	5.47	0.13	56.60	10.93
FG	0.00	41.22	16.32	2	7	4.75	0.00	74.47	8.01	10.00	41.22	18.13	2	7	4.68	0.40	84.47	19.99
FS	0.00	84.33	23.99	0	17	4.71	0.00	84.47	8.54	7.04	76.17	36.77	0	12	3.50	0.13	55.25	9.92

Table 6.5b Secondary results of SURFACE II GRAPHICS (OCTANT)

quadrants; in addition, within each octant, two points occur within the QUADRANT range and two points occur within the OCTANT range but outwith the QUADRANT range, the other quadrants containing no points. In this latter case, the OCTANT search will be successful and the QUADRANT search will fail. While it is not possible to compute an authoritative correlation due to the preponderance of zero values in %BAD, nevertheless a strong rank correlation exists between the non-zero %BAD values and NNS. Similarly, a significant rank correlation exists between the individually averaged ORIGINAL MIN and MAX magnitudes and the NNS. These are similar to the DEFAULT and QUADRANT results. More interestingly, the individual OCTANT ORIGINAL MIN and MAX values adhere closely to the relevant QUADRANT values. Indeed, the overall mean magnitude of the ORIGINAL MIN/MAX statistics differs by only 0.06 (QUADRANT - 15.84 OCTANT - 15.78), while differing noticeably from the DEFAULT equivalent (17.33). Additionally, the IN and OUT, QUADRANT and OCTANT statistics behave more similarly - particularly the NEAR-MEAN statistics. The IN-MEAN is greater for the OCTANT and QUADRANT values, while OUT-MEAN is smaller. These factors stress the benefits of applying either form of directional search. Nodes based on distant neighbouring points can have acceptable interpolated values, due to obtaining a more representative balanced distribution for the interpolation.

The accuracy descriptors show similar effects to QUADRANT. STANDARD ERROR has a highly significant (0.89) rank correlation, and ABS MEAN (0.84) and SURFACE CORRELATION (0.82) have significant rank correlations with NNS. All have

important secondary correlations with two-dimensional autocorrelation (mean - 0.65). While STANDARD ERROR has a similar global mean (3.05) to DEFAULT (3.09) and QUADRANT (2.95), there are distinctly different responses for the three surfaces where autocorrelation and NNS are concerned. In the case of the smooth LIAN surface the accuracy descriptors have generally improved for OCTANT. For the rougher INCH surface, accuracy - although improved over using the DEFAULT search - is lower than for QUADRANT. Results for the FORV data display a further decrease in accuracy over even the QUADRANT searching. Manifestly, the OCTANT search involves more directionality than the QUADRANT search and therefore, in general, even more distant points will be selected. The significance of this in relation to the rapid decay in autocorrelation in FORV data has already been stressed for QUADRANT searching and the effect will be even more acute for OCTANT searching.

The difference between LIAN and INCH is more difficult to explain though the major causes are easily assigned. LIAN is a smooth two-slope linear surface, and therefore so long as a good, balanced distribution of data points is obtained in the search (which is the reason for using a directional search), the weighting function should produce good results and as greater directionality is applied (DEFAULT to QUADRANT to OCTANT searching), accuracy gradually improves. The INCH data however, describe a much rougher surface, with many sharp breaks of slope; therefore a weighted average between a few distant points cannot pick up local innuendos in the data. Clearly, the OCTANT search has forced the averaging function

to use too-distant points. In this situation, a little directionality is desirable, but not too much.

The error-related correlation coefficients show a continuation of the QUADRANT trend. Both ERROR/NEAR and ERROR/NONN show minor increases in their mean value, once again stressing the importance of having near neighbouring data points round the grid node to be interpolated. Mean ERROR/GRAD (0.124) displays a very highly significant drop in correlation against QUADRANT ERROR/GRAD (0.193), which displayed only a significant drop against DEFAULT (0.232). This stresses the fact that the data points found are more distant as the directionality of the search increases, and therefore the local gradient in the immediate vicinity of the grid node is of even less importance. Of more importance are variations in the local gradients in the vicinity of the data points and in the direction of the quadrants and octants.

The other, precision and systematic error statistics behave similarly. The mean magnitude TRIMMED MIN and MAX value decreases to 7.89 (DEFAULT is 8.26, QUADRANT is 7.83) suggesting a minor increase in precision. However, on closer examination, the FORV data sets increase and the LIAN data sets decrease, while the INCH data sets produce variable results. This is in accordance with the mechanisms postulated in the previous discussion. Additionally, mean KURTOSIS shows a continuous decrease (DEFAULT is 3,48, QUADRANT is 3.39 and OCTANT is 3.27). While this global mean hides the increase for the high NNS GRID data, in general the distributions are becoming less peaked, with low error increasing and high error

decreasing, as discussed for QUADRANT. Systematic error shows a minor positive swing for both the mean MEAN (DEFAULT - -0.14, QUADRANT - -0.15 and OCTANT - -0.10) and the mean MEDIAN (DEFAULT - -0.14, QUADRANT - -0.12 and OCTANT - -0.07), however, this may be subdivided into a minor negative swing for INCH, and a minor positive swing for LIAN and FORV. Interestingly, this general mean MEDIAN trend corresponds with the altitude skewnesses generated by GEOD in the surface characteristic analysis. Finally, SKEWNESS once again shows little change, although minor fluctuations occur.

Unfortunately, purely statistical summaries do not fully describe relative performances of the DEFAULT, QUADRANT and OCTANT searches which are better summarised in Figures 6.1, 6.4 and 6.5. While INCH-GRID (Figure 6.4) shows little differences between the three search strategies as a result of higher autocorrelation, INCH-BREAKLINE and FORV-GRID demonstrate the smoothing effect of the OCTANT search. In general, the area enclosed by high value isarithms decreases, while that enclosed by low value isarithms increases. Because a direct weighted average is used the range of the interpolated grid nodes is less than the range of the data points used to derive the interpolated grid. As the width of the search segment diminishes and consequently, as more distant (and therefore probably more variable) data points are used, the effect of smoothing on the range of the interpolated grid increases. This removes some isarithm levels completely, but also increases the number of small circular 'adopted isarithms' - FORV-GRID DEFAULT (20 occurrences), FORV-GRID QUADRANT (30 occurrences) and FORV-GRID OCTANT (40

occurrences). Since surface autocorrelation provides a descriptor of the decay of data point similarity with increasing distance, the autocorrelation functions may be considered a good pointer to the expected smoothing, and thus the likely number of adopted isarithms from a particular data set.

6.2.1.4 VRADIUS

Unlike the fixed number of points found by the DEFAULT search procedure, the VRADIUS (Variable RADIUS) search procedure may find any number within a minimum and maximum specified (8 and 12 for all data sets). These may be found within a minimum radius from the grid node, or alternatively the radius is iteratively expanded over a specified number of equal steps to the maximum radius. For all data sets, 6 iterations could be used, with minimum and maximum radii defined as 120m and 1000m (INCH), 60m and 400m (LIAN) and 50m and 500m (FORV). These parameters were chosen by considering the parameters of previous searches in conjunction with the manual (Sampson, 1978).

VRADIUS is essentially an improved, more flexible form of the DEFAULT search which is guaranteed to produce a more successful search of the nearest (and therefore more likely similar) points. This has two effects. The points found should be closer to the grid node but without the use of the directionality constraint of QUADRANT or OCTANT. Thus, where data sets have a high NNS, the final accuracy may be improved on DEFAULT searching but may possibly be poorer in relation to

those from QUADRANT or OCTANT searching. Where high terrain variability exists (e.g. in the FORV data), the accuracy should be better than with QUADRANT - which includes more distant points. With the parameters used, the only difference between DEFAULT and VRADIUS is that the latter may include an additional one, two, or three points (which will be the most distant points); the other points used in the estimation process will be identical. Thus results should thus be virtually identical with DEFAULT, perhaps producing slightly inferior results in areas of high terrain variability and low areal correlation.

It is apparent from considering the VRADIUS interpolation statistics (see Table 6.6) that little difference exists between using VRADIUS and DEFAULT. The major difference is in the consistency as expressed by %BAD and the ORIGINAL MIN/MAX. Again the individual mean ORIGINAL MIN/MAX values are significantly rank-correlated with NNS. The increase in the maximum search radius has resulted in very few failed searches, restricted to INCH-BREAKLINE and -RIVER - i.e. the data sets with the lowest NNSs in INCH. Unfortunately, the mean ORIGINAL MIN/MAX magnitude (17.40 against DEFAULT - 17.33) does not reflect this improvement. This illustrates the danger of relying too heavily on means since, on closer examination, the general trend is for decreased MIN/MAX magnitude which has been swamped by a dramatic increase in the magnitude of the LIAN-RIVER ORIGINAL MIN and MAX values.

These extreme errors are a direct result of relying on a nearest neighbour type search which may go uni-directional in

DS	%BAD	ORIGINAL		TRIMMED		TRIMMED ERROR				TRIMMED ABSOLUTE ERROR			CORRELATIONS			
		MIN	MAX	MIN	MAX	MEAN	MEDIAN	SKEWNESS	KURTOSIS	MEAN	MEDIAN	STANDARD	SURFACE	NEAR	NONN	GRADIENT
IB	0.21	(-33.77	38.88)	-14.89	14.92	1.01	0.80	0.06	3.22	4.27	3.15	5.61	0.992	0.298	-0.213	0.216
IC	0.00	-18.86	26.67	-10.52	10.56	0.08	0.08	0.10	2.76	3.31	2.73	4.18	0.995	0.127	0.014	0.108
IG	0.00	-19.72	21.27	-8.96	8.97	0.05	-0.03	-0.04	3.18	2.39	1.84	3.22	0.997	0.523	-0.454	0.252
IR	1.46	(-69.10	32.11)	-27.46	26.88	-5.60	-4.84	0.16	4.02	7.89	6.10	10.25	0.983	0.488	-0.376	0.300
IS	0.00	-21.39	38.74	-13.26	13.22	0.60	0.79	-0.23	3.44	3.57	2.72	4.68	0.994	0.211	-0.052	0.308
LB	0.00	-7.47	7.43	-4.47	4.40	0.22	0.25	-0.15	2.81	1.40	1.15	1.77	0.996	0.231	-0.160	0.122
LC	0.00	-7.35	7.31	-4.27	4.39	0.31	0.05	0.37	2.60	1.39	1.14	1.76	0.997	0.305	-0.253	0.196
LG	0.00	-7.84	4.50	-2.55	2.40	0.01	-0.14	0.01	3.11	0.71	0.56	0.93	0.999	0.454	-0.074	0.290
LR	0.00	-20.25	20.64	-11.85	12.10	-0.71	-0.75	0.66	5.89	2.32	1.38	3.44	0.989	0.585	-0.470	0.500
LS	0.00	-11.48	8.77	-3.96	3.97	0.03	-0.03	0.12	3.14	1.10	0.88	1.43	0.998	0.072	0.060	0.330
FB	0.00	-8.43	8.64	-4.93	4.98	0.68	0.67	-0.15	3.71	1.39	1.05	1.81	0.947	0.363	-0.322	0.107
FC	0.00	-7.02	6.29	-3.93	3.97	0.77	0.87	-0.55	3.68	1.31	1.14	1.60	0.964	0.180	-0.008	0.031
FG	0.00	-10.35	4.61	-2.77	2.72	0.24	0.27	-0.44	3.51	0.75	0.60	0.98	0.983	0.362	-0.018	0.138
FS	0.00	-8.31	9.72	-4.38	4.47	0.53	0.34	0.44	3.64	1.05	0.73	1.44	0.967	0.287	-0.143	0.281

Table 6.6a Primary results of SURFACE II GRAPHICS (VRADIUS)

DS	IN-NEAR			IN-NONN			IN-GRAD			OUT-NEAR			OUT-NONN			OUT-GRAD		
	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN
IB	0.95	174.83	50.17	0	14	3.81	0.12	83.55	18.91	18.83	144.69	77.47	0	9	1.93	3.38	83.00	28.00
IC	0.83	182.47	43.46	0	10	3.32	0.12	83.55	19.00	12.27	157.25	61.17	0	9	3.25	0.17	83.00	26.67
IG	0.00	42.42	32.19	4	6	4.21	0.17	83.00	18.82	30.00	42.42	36.94	4	5	4.04	0.12	83.55	30.10
IR	0.91	306.41	70.55	0	20	3.46	0.12	83.55	18.71	42.42	305.62	164.72	0	5	0.24	3.71	78.75	31.82
IS	0.31	148.00	45.61	0	12	3.36	0.12	83.55	18.75	21.83	150.02	69.25	0	6	2.36	2.90	78.75	31.47
LB	0.31	165.00	39.16	0	23	3.23	0.00	41.50	10.15	9.48	145.94	55.78	0	8	1.27	0.76	29.74	14.96
LC	0.00	133.99	27.08	0	11	3.36	0.00	41.50	10.67	9.02	117.52	54.22	0	8	1.15	0.50	32.27	5.31
LG	0.00	42.42	19.58	1	6	3.35	0.00	41.50	10.02	15.00	42.42	24.19	1	5	2.96	2.69	39.77	17.77
LR	0.31	226.99	52.96	0	19	3.42	0.00	41.50	10.04	46.94	221.52	149.13	0	1	0.01	6.17	33.14	17.30
LS	0.21	112.92	27.41	0	11	3.05	0.00	41.50	10.01	4.86	63.55	33.00	0	9	2.89	0.74	33.14	17.86
FB	0.21	129.38	27.08	0	18	4.82	0.00	84.47	8.61	9.29	129.47	72.11	0	12	0.90	0.13	56.39	8.35
FC	0.40	70.00	21.30	0	17	4.56	0.00	84.47	8.50	5.47	47.27	23.32	1	15	6.09	0.13	55.25	10.59
FG	0.00	41.22	16.27	2	7	4.76	0.00	84.47	7.97	10.00	41.22	18.83	2	7	4.60	0.91	74.47	20.78
FS	0.00	84.33	24.27	0	17	4.63	0.00	84.47	8.23	7.42	67.88	31.22	0	12	4.85	0.56	61.88	15.57

Table 6.6b Secondary results of SURFACE II GRAPHICS (VRADIUS)

a low NNS data set. The 'UNI-DIRECTIONAL EFFECT' is described in one dimension in Figure 6.7. Where a linear slope exists such as in LIAN, and the grid is being interpolated by a weighted average interpolation method from a low NNS data set, the search will fail (in the example) if a DEFAULT search is used with a 100m maximum search radius. If a directional search (QUADRANT or OCTANT) is used with an increased radius (150m), then a more balanced distribution in the plane, and in height, will be obtained (perhaps the nearest four points on either side), producing a reasonable average (c). However, if a non-directional increased radius (150m) search (VRADIUS) is used then there is a danger that only the eight right-most points will be used, producing a surface-unrepresentative average (d), with high positive error. If the slope continues past b. and the same situation occurs, then a high negative error will result at point b.

This contention is supported by comparing the IN and OUT, NEAR and NONN values for the DEFAULT and VRADIUS searches. In general, nodes with more distant nearest neighbours (IN-NEAR MAX increased from 103.08 to 226.99) and less close neighbours (IN-NONN MEAN decreased from 4.44 to 3.42) are successfully interpolated by VRADIUS.

A similar pattern emerges when the precision statistics are considered. Generally, TRIMMED MINs and MAXs are smaller, although LIAN-RIVER is distinctly larger. This is not just a problem of extreme errors within LIAN-RIVER, but is inherent throughout the error distribution as a result of the change in gradient across the model. In the UNI-DIRECTIONAL EFFECT, a

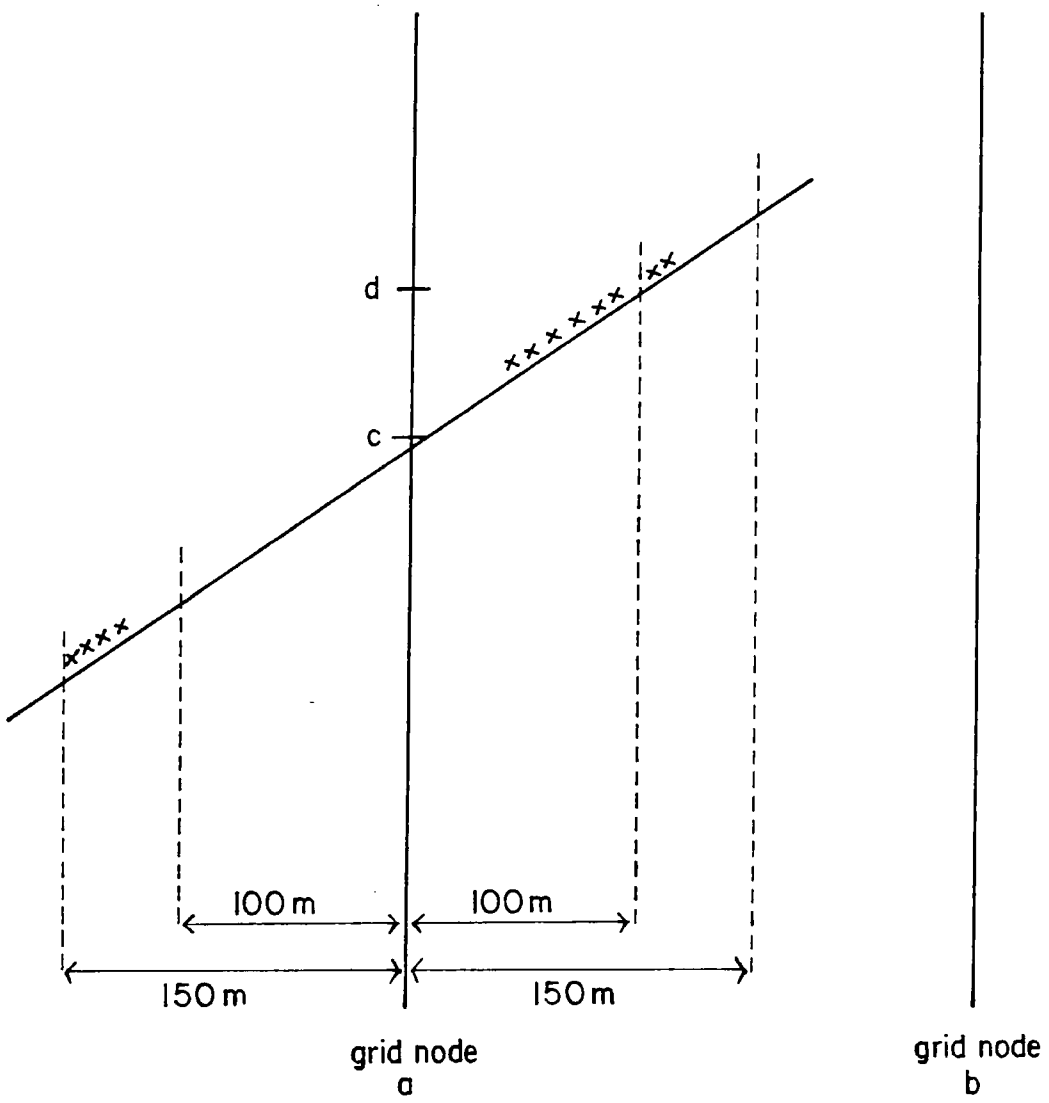


Figure 6.7 The Uni-directional effect, shown in cross section.

flat area will have less error than a steep area, and a wide variety of situations must exist in the LIAN-RIVER data, since the 26.96 %BAD value for DEFAULT searching has been completely removed for VRADIUS.

Mean KURTOSIS similarly is stable (3.48 for DEFAULT and VRADIUS), although the general trend is for decreasing KURTOSIS (as for QUADRANT and OCTANT). The KURTOSIS values from LIAN-RIVER data have increased dramatically, although this may be a direct result of the high increase in SKEWNESS in combination with the increased TRIMMED range.

The accuracy and systematic error statistics show a marked similarity with NEAR - mean STANDARD ERROR (VRADIUS - 3.08 and DEFAULT - 3.09) and mean MEAN (VRADIUS - -0.13 and DEFAULT - -0.14). STANDARD ERROR (0.91) and SURFACE CORRELATION (0.91) have very highly significant , and ABS MEAN has a highly significant rank correlation with NNS. Importantly, higher secondary correlations exist with areal autocorrelation - ABS MEAN (0.82), and STANDARD ERROR and SURFACE CORRELATION (0.71). These stress the importance of the wider environment in this search method. However, the MEANs are very similar, supporting the concept that, by not applying any directionality, the NEAR and VRADIUS are inherently similar and areas of positive and negative error will similarly balance each other.

The error-related correlations support the previous conclusions. While the correlations of the low NNS data sets have increased more than the average, the general trend is for increased strength of the ERROR/NEAR and ERROR/NONN

correlations, while decreasing the strength of the ERROR/GRAD correlations. The contributing data points to the interpolation are generally more distant for this method, and therefore local grid node gradients are not nearly as important as is the data distribution around the grid node, especially as a result of increased number of nodes successfully interpolated. It is noteworthy that the VRADIUS correlations are similar to the QUADRANT and, to a lesser extent, the OCTANT correlations. These other methods similarly involve using more distant points and therefore would be expected to be the same.

The similarity between DEFAULT and VRADIUS is emphasised in Figures 6.2 and 6.3. Use of INCH-GRID data produces identical results to the DEFAULT search while, within FORV-GRID data, VRADIUS produces marginally superior results in the middle range of isarithms and marginally poorer results at the extremes. If adopted isarithms are a measure of accuracy, then VRADIUS has a lower accuracy (27 occurrences compared with 20 for DEFAULT). The improved performances of VRADIUS over QUADRANT are emphasised in the results from the FORV-BREAKLINE data (Figure 6.6). While considerable error still exists (top-right-of-centre), VRADIUS maintains a better general fit especially at the higher and lower isarithm levels.

6.2.1.5 Weighting

The effect of utilising different weighting functions within pointwise interpolation was evaluated using the SCATTER data sets for each model within the various weighting options

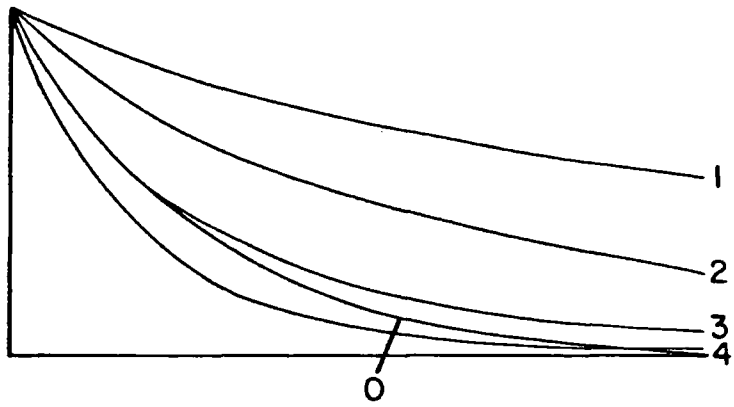
in SURFACE II GRAPHICS (as outlined in 2.5.3 and illustrated in Figure 6.8). This replicated work performed earlier by Davis, who came to the conclusion that

> "weighting functions which drop off slowly have the lowest bias but the highest RMS error" (Davis, 1975, 357).

That this is perhaps an over-generalisation may be seen from Table 6.7. The weighting function with the lowest bias (MEAN) and highest RMS error (STANDARD ERROR) is usually, but not always, W1 (1/d). In the LIAN-SCATTER statistics, W3 has the lowest magnitude bias, but W0 has the highest STANDARD ERROR - both are based on functions with a much faster decay than W1. Similarly, the FORV-SCATTER data set has the lowest STANDARD ERROR for W1. The Davis theory may be completely rejected when the other functions are considered in order of steepness of decay. MEAN error has a mean rank correlation of 0.23 (INCH - 1.0, LIAN - -0.7 and FORV - 0.4), and STANDARD ERROR has a mean rank correlation of 0.03 (INCH - 0.1, LIAN - 0.3 and FORV - -0.3) against function steepness.

A relationship may, however, be seen to exist between data/surface characteristics and optimum weighting function. The FORV data produces the highest NNS, suggesting a more regular distribution of points round each node and a greater consistency in maximum search radius. INCH and, to a greater extent, LIAN data produce lower NNSs and will therefore have lower consistency of point distribution and thus require greater search radii around each node, with greater variance within each sought area. Thus FORV should utilise a slower decay function and LIAN a steeper one, with INCH using an intermediate value. Vector dispersion statistics confirm this

Figure 6.8



0 $w = (1 - \frac{d}{1{\cdot}1 \times d_{max.}})^2 / (\frac{d}{1{\cdot}1 \times d_{max.}})^2$

1 $w = 1 / d$

2 $w = 1 / d^2$

3 $w = 1 / d^4$

4 $w = 1 / d^6$

Generalised plot of weighting functions available in SURFACE II GRAPHICS.

(After Sampson, 1978, 95.)

DS	%BAD	ORIGINAL		TRIMMED		TRIMMED ERROR				TRIMMED ABSOLUTE ERROR			CORRELATIONS			
		MIN	MAX	MIN	MAX	MEAN	MEDIAN	SKEWNESS	KURTOSIS	MEAN	MEDIAN	STANDARD	SURFACE	NEAR	NONN	GRADIENT
I0	0.00	-20.89	38.74	-13.36	13.35	0.62	0.82	-0.22	3.42	3.65	2.81	4.80	0.994	0.224	-0.060	0.312
I1	0.00	-29.14	35.66	-15.26	15.19	0.32	0.52	-0.05	3.21	4.22	3.26	5.44	0.993	0.109	-0.004	0.188
I2	0.00	-26.83	33.16	-13.40	13.65	0.47	0.60	-0.15	3.39	3.55	2.77	4.64	0.995	0.184	-0.003	0.234
I3	0.00	-23.05	35.32	-13.55	13.46	0.61	0.83	-0.21	3.32	3.71	2.80	4.85	0.994	0.190	-0.052	0.340
I4	0.00	-22.30	39.55	-14.15	14.11	0.65	0.80	-0.20	3.18	4.05	3.14	5.27	0.993	0.213	-0.060	0.358
L0	0.00	-12.18	8.97	-4.15	4.23	0.00	-0.08	0.18	3.36	1.14	0.89	1.48	0.998	0.068	0.069	0.345
L1	0.00	-8.14	8.60	-3.76	3.77	0.37	0.35	-0.09	2.72	1.22	1.05	1.54	0.998	0.006	0.073	0.164
L2	0.00	-8.97	8.52	-3.35	3.36	0.21	0.17	-0.03	2.72	1.06	0.90	1.34	0.998	0.076	0.062	0.237
L3	0.00	-10.29	8.98	-3.97	3.97	0.02	-0.03	0.12	3.13	1.11	0.89	1.43	0.998	0.056	0.064	0.321
L4	0.00	-11.19	9.76	-4.52	4.52	-0.06	-0.12	0.10	3.43	1.22	0.93	1.59	0.997	0.053	0.074	0.366
F0	0.00	-8.89	11.02	-4.93	5.13	0.79	0.54	0.43	3.57	1.27	0.88	1.75	0.954	0.266	-0.182	0.189
F1	0.00	-8.35	10.31	-4.14	4.23	0.50	0.32	0.40	3.56	1.00	0.68	1.36	0.970	0.258	-0.111	0.287
F2	0.00	-10.72	7.25	-5.04	5.07	0.94	1.02	-0.44	3.31	1.72	1.46	2.14	0.927	0.337	-0.292	-0.020
F3	0.00	-9.42	8.92	-4.89	4.94	0.76	0.66	-0.13	3.52	1.40	1.05	1.84	0.948	0.304	-0.248	0.077
F4	0.00	-9.39	9.67	-5.13	5.21	0.89	0.67	0.21	3.48	1.42	1.01	1.89	0.947	0.300	-0.222	0.141

Table 6.7a Primary results of SURFACE II GRAPHICS (weighting options)

DS	IN-NEAR			IN-NONN			IN-GRAD			OUT-NEAR			OUT-NONN			OUT-GRAD		
	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN
I0	0.31	148.00	45.63	0	12	3.36	0.12	83.55	18.69	21.83	150.02	68.69	0	6	2.39	5.32	78.75	32.46
I1	0.31	148.00	45.77	0	12	3.36	0.12	83.55	18.91	12.89	150.02	66.22	0	7	2.31	2.90	67.19	28.47
I2	0.31	148.00	45.53	0	12	3.38	0.12	83.55	18.85	26.50	150.02	70.53	0	6	2.22	2.90	77.96	29.44
I3	0.31	148.00	45.66	0	12	3.36	0.12	83.55	18.74	21.83	150.02	68.08	0	7	2.47	2.90	78.75	31.69
I4	0.31	150.02	45.77	0	12	3.36	0.12	83.55	18.69	21.83	131.77	65.96	0	7	2.48	5.32	78.75	32.58
L0	0.21	105.08	27.19	0	11	3.06	0.00	41.50	10.02	10.34	112.92	36.86	0	9	2.60	0.74	33.14	17.58
L1	0.21	112.92	27.58	0	11	3.06	0.00	41.50	10.18	5.27	60.58	29.61	0	7	2.59	1.11	32.27	14.47
L2	0.21	112.92	27.44	0	11	3.06	0.00	41.50	10.13	4.86	63.55	32.46	0	9	2.82	1.11	33.14	15.68
L3	0.21	112.92	27.50	0	11	3.05	0.00	41.50	9.97	4.86	63.55	31.33	0	9	2.96	1.25	33.14	18.47
L4	0.21	112.92	27.39	0	11	3.06	0.00	41.50	9.97	9.26	63.55	33.17	0	8	2.72	0.74	33.14	18.67
F0	0.00	80.61	24.08	0	17	4.68	0.00	84.47	8.50	8.47	84.33	34.86	0	12	3.96	0.13	55.25	10.68
F1	0.00	84.33	24.33	0	17	4.63	0.00	84.47	8.22	7.42	62.22	30.21	0	12	5.04	0.56	61.88	15.82
F2	0.00	84.33	24.05	0	17	4.73	0.00	84.47	8.61	3.52	72.11	35.61	0	12	2.84	0.13	56.60	8.40
F3	0.00	84.33	24.07	0	17	4.69	0.00	84.47	8.55	8.46	76.80	35.25	0	12	3.60	0.13	55.25	9.64
F4	0.00	84.33	24.00	0	17	4.71	0.00	84.47	8.54	7.04	76.17	36.77	0	12	3.50	0.13	55.25	9.92

Table 6.7b Secondary results of SURFACE II GRAPHICS (weighting options)

theory since FORV is considered smoother than LIAN and INCH.

On the other hand, autocorrelation coefficients suggest that LIAN should employ the slower decay, and FORV employ a more rapidly decaying function. LIAN data produces a slower autocorrelation decay, suggesting a more highly autocorrelated surface. However, if the distribution is initially even, the search should not include more distant points and thus the autocorrelation function should not be considered as important in this case.

The error-related correlations are of more help. Mean ERROR/GRAD correlations of 0.135 (FORV) and 0.287 (LIAN and INCH) and mean ERROR/NEAR correlations of 0.293 (FORV), 0.184 (INCH) and 0.052 (LIAN) are recorded. This suggests that, for INCH and LIAN data, the local gradient around the interpolated node is crucial. Since 'near points are similar and distant points differ', this suggests that a slightly faster decay function should be generally used, since where steep gradients occur, distant points will differ more greatly. In contrast, the lesser importance of nearness of data points for FORV-SCATTER supports the use of a slower decay function - where the closest point is relatively distant, it should have a similar weighting to the other points selected.

Thus, in conclusion, where NNSs are similar, autocorrelation coefficients are useful indicators of the best weighting functions; in general, higher NNSs or lower vector dispersion statistics suggest the requirement for slower decay weighting functions. An interesting feature of this set of statistics is the lack of variance among the same statistics

within each data set. While this is less than that found as a result of varying the data set within a specific method, or varying the package with the same data set, it is greater than that found by varying the search method for any particular data set.

6.2.2 GPCP

Grid estimation may be performed by GPCP by utilising two types of function. The 'TREND' routine involves global interpolation and will be examined later. The default method involves a two-stage, pointwise weighted average, as described in 2.5.1. This was accessed in its most basic default form with the weighting factor w=1/(1-r), the two-stage, interpolation process involving the averaging of local linear surfaces constrained to pass through the data points and search routines set to find the eight nearest neighbours (see Table 6.8). Excluding the minor effects of a different weighting function, this may thus be used to compare the direct data point averaging technique of SURFACE II GRAPHICS - DEFAULT with the two-stage surface averaging technique of GPCP. Unlike SURFACE II GRAPHICS, GPCP evaluated every grid node and therefore the %BAD points refer only to those found outside the respective tolerances by the 'MATANN' program. These extreme values were all located at internal nodes and therefore they cannot be related to 'EDGE-EFFECT' (distortion at the edges of the data areas).

The effects of the superior search method within GPCP are visible within the consistency statistics. The %BAD statistics

DS	%BAD	ORIGINAL		TRIMMED		TRIMMED ERROR				TRIMMED ABSOLUTE ERROR			CORRELATIONS			
		MIN	MAX	MIN	MAX	MEAN	MEDIAN	SKEWNESS	KURTOSIS	MEAN	MEDIAN	STANDARD	SURFACE	NEAR	NONN	GRADIENT
IB	0.25	-66.41	42.13	-16.19	15.93	0.12	0.27	-0.15	3.80	3.90	2.82	5.27	0.993	0.337	-0.241	0.058
IC	0.00	-29.27	22.10	-10.51	10.51	-0.33	-0.31	0.05	3.00	3.05	2.42	3.90	0.996	0.200	-0.012	-0.023
IG	0.00	-14.77	17.55	-7.17	7.18	-0.18	-0.77	-0.01	3.25	1.90	1.50	2.56	0.998	0.545	-0.453	0.117
IR	6.96	-97.75	78.00	-48.60	45.16	-6.56	-4.22	-0.49	5.15	9.72	5.92	14.14	0.966	0.505	-0.371	0.277
IS	0.21	-48.07	30.47	-12.25	12.38	0.80	0.67	-0.10	3.38	3.22	2.50	4.19	0.996	0.227	-0.016	0.159
LB	0.00	-20.00	17.92	-6.47	6.34	-0.90	-0.75	0.01	3.77	1.69	1.27	2.27	0.995	0.351	-0.251	-0.088
LC	0.00	-11.22	7.68	-3.75	3.77	-0.31	-0.45	0.50	3.11	1.17	0.98	1.46	0.997	0.274	-0.213	-0.191
LG	0.00	-7.14	5.68	-2.23	2.22	-0.33	-0.14	0.04	3.38	0.58	0.46	0.77	0.999	0.476	-0.085	0.142
LR	0.00	-18.71	12.81	-8.31	7.73	-1.10	-0.87	-0.25	4.02	2.01	1.38	2.76	0.994	0.468	-0.352	0.347
LS	0.00	-7.14	10.67	-3.77	3.75	-0.38	-0.40	0.29	3.32	1.09	0.88	1.39	0.998	0.198	-0.081	0.063
FB	0.00	-9.17	25.88	-8.43	8.39	0.21	-0.08	0.64	5.42	1.47	0.97	2.15	0.933	0.433	-0.328	0.196
FC	0.00	-7.89	12.32	-3.00	3.02	0.06	0.03	-0.02	3.47	0.78	0.60	1.02	0.981	0.130	-0.095	0.086
FG	0.00	-10.31	6.02	-2.32	2.29	0.11	0.07	0.09	4.34	0.49	0.34	0.68	0.992	0.343	-0.073	0.244
FS	0.00	-11.88	16.82	-5.23	5.40	0.05	-0.08	0.54	5.00	1.04	0.70	1.48	0.962	0.240	-0.046	0.211

Table 6.8a Primary results of GPCP

DS	IN-NEAR			IN-NONN			IN-GRAD			OUT-NEAR			OUT-NONN			OUT-GRAD		
	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN
IB	0.95	168.94	49.11	0	14	3.86	0.12	83.55	19.33	18.83	174.83	98.99	0	8	1.06	0.89	77.96	20.44
IC	0.83	182.47	42.89	0	10	3.35	0.12	83.55	19.41	12.27	173.91	71.80	0	9	2.85	0.17	83.00	18.83
IG	0.00	42.42	32.19	4	6	4.21	0.17	83.00	19.00	30.00	42.42	37.03	4	5	4.02	0.12	83.55	26.72
IR	0.91	306.41	70.66	0	20	3.47	0.12	83.55	19.00	66.78	297.06	165.82	0	2	0.10	2.64	67.19	26.50
IS	0.31	150.02	45.52	0	12	3.35	0.12	83.55	19.05	14.97	140.80	70.86	0	9	2.60	3.36	78.75	25.80
LB	0.31	165.00	37.22	0	23	3.27	0.00	41.50	10.51	19.72	160.88	92.94	0	10	0.50	0.56	30.24	8.30
LC	0.00	133.99	27.33	0	11	3.35	0.00	41.50	10.60	7.11	123.78	49.25	0	8	1.63	0.35	27.72	6.73
LG	0.00	42.42	19.52	1	6	3.34	0.00	40.00	10.06	15.00	42.42	25.35	1	4	2.98	2.23	41.50	16.91
LR	0.31	226.99	53.77	0	19	3.39	0.00	41.50	10.14	18.88	225.50	133.61	0	5	0.36	2.90	35.99	15.18
LS	0.21	112.92	27.27	0	11	3.05	0.00	41.50	10.27	4.86	86.11	35.58	0	9	3.01	0.56	36.63	12.97
FB	0.21	129.38	27.36	0	18	4.81	0.00	84.47	8.57	12.85	129.47	67.11	0	10	1.22	0.13	47.52	9.11
FC	0.40	70.00	21.08	0	16	4.55	0.00	74.47	8.35	3.77	65.19	27.66	0	17	6.48	0.28	84.47	13.31
FG	0.00	41.22	16.22	2	7	4.77	0.00	84.47	7.89	10.00	40.00	20.00	2	7	4.47	1.07	74.47	22.11
FS	0.00	84.33	23.97	0	17	4.67	0.00	84.47	8.31	7.42	78.64	37.11	0	12	4.13	0.56	61.88	14.34

Table 6.8b Secondary results of GPCP

are greatly reduced throughout, although the only non-zero values are strongly correlated with the NNSs. This is reflected in the ORIGINAL MIN and MAX values which have generally increased (the mean magnitude has risen from DEFAULT - 17.33 to GPCP - 23.78). This might be expected since more nodes are interpolated where the neighbours are extremely distant, however it is not necessarily automatic. Very few nodes were unsuccessfully searched in OCTANT and QUADRANT and yet the mean magnitude ORIGINAL MIN/MAX decreased to 15.78 and 15.84. Indeed, in specific situations where OCTANT and GPCP were fully successful, the magnitudes of the MIN and MAX values have increased dramatically (e.g. for INCH-RIVER data, OCTANT - -6.71 and 19.37, GPCP - -97.75 and 78.00). There are, however, a few situations where the magnitudes of the MIN and MAX errors have decreased, although this appears random and not necessarily attributed to any characteristic of the data (e.g. with INCH-GRID and LIAN-SCATTER data).

A more interesting feature of the consistency descriptors is the decrease in balance between the magnitude of the ORIGINAL MIN and MAX errors. In general, there is a stronger negative bias to the error, although individual cases have received dramatic changes (e.g. results from INCH-BREAKLINE and INCH-SCATTER data show increasing negative bias, whereas those from LIAN-SCATTER and FORV-CONTOUR data show increasing positive bias). This suggests that a stable process has been replaced with a less stable process, although it should be remembered that there are many nodes which have not been interpolated by DEFAULT.

From a preliminary, global examination of the TRIMMED MIN and MAX values, it appears that GPCP has an inferior precision with respect to DEFAULT (mean magnitudes of 9.73 against 8.26). On closer examination, this may be broken down into a two-way effect directly related to NNS. In all models, the highest NNS data sets have increased their precision, while the low NNS data sets (INCH-RIVER, LIAN-RIVER, LIAN-BREAKLINE, FORV-BREAKLINE and FORV-SCATTER) have decreased their precision. Indeed, if INCH-RIVER were removed, the mean magnitude for GPCP would only be 8.29.

The other precision descriptor, KURTOSIS, contradicts this similarity and importantly suggests that GPCP has a much higher precision. Only results from INCH-SCATTER, LIAN-RIVER, LIAN-SCATTER and FORV-CONTOUR data show minor decreases in KURTOSIS which, in general, rises from a mean of 3.22 to 3.89. This suggests that, within GPCP, large errors become greater, while low errors become closer to the mean error. That is, where good control exists and low error is found, error should become closer to the systematic error of the method, whereas conditions which create large errors within weighted average methods will be magnified by GPCP.

Associated with the increase in precision is the increased negative systematic error of GPCP. The DEFAULT mean MEAN value (-0.15) has increased to -0.63 for GPCP. With the exclusion of INCH-SCATTER data, all MEANs show an increase in negative bias. However, as with DEFAULT, use of RIVER and CONTOUR data sets produce distinct negative bias, GRID data produces low bias and SCATTER and BREAKLINE data sets produce more variable

results. The general characteristics of the data sets have still been preserved.

A very highly significant aspect of GPCP interpolation is the very strong 0.96 rank correlation between all the accuracy descriptors and NNS. There is also no significant secondary correlation with any other data/surface characteristic. Moderate increases in accuracy occur with the use of data sets with both a high accuracy using weighted average interpolation and a high NNS value - notably all GRID and CONTOUR and SCATTER data sets. Those data sets with a lower weighted average accuracy and poorer NNS perform less well, several displaying a marked decrease in accuracy (such as the INCH-RIVER, LIAN-RIVER, LIAN-BREAKLINE and FORV-BREAKLINE data sets).

The importance of near neighbours to high precision, accuracy and consistency is stressed by the error-related correlations. In general, the mean ERROR/NEAR (0.338) and ERROR/NONN (-0.187) are significantly stronger, and the ERROR/GRAD (0.114) is significantly weaker than are the respective error-related correlations for the DEFAULT (NEAR) SURFACE II GRAPHICS option. Similarly, the GPCP IN-GRAD MEANs and OUT-GRAD means are much more similar and consistent than for DEFAULT, underlining the lack of importance of local gradient in this method. These changes are a direct response to the fact that interpolated values from local surfaces, and not actual data values, are averaged. In the local surface evaluation, gradient at the data point will no doubt be of considerable importance. However, when interpolated values

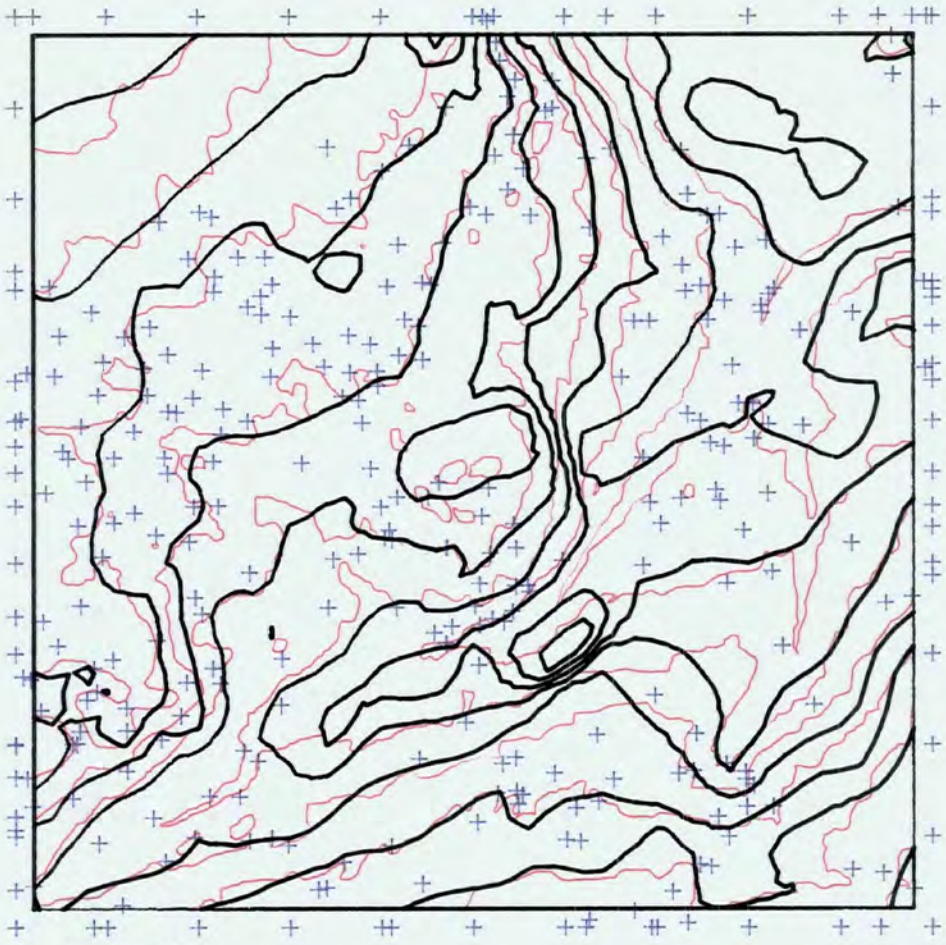
from these surfaces areaveraged,it is only important that the node/data point distance be minimised, since the smaller the distance, the more reasonable it is to assume that the local trend at the data points is also that at the grid node.

The superior performance of this interpolator is demonstrated by considering the changing values of the IN-NEAR and OUT-NEAR statistics between DEFAULT and GPCP. While most IN-NEAR MAX values and OUT-NEAR MIN values have increased for GPCP, almost without exception, all the IN-NEAR and OUT-NEAR MEAN values have increased. This suggests that, with more distant data values than for the DEFAULT option of SURFACE II GRAPHICS, GPCP can produce acceptable results, while the data points must be very distant to cause severe problems.

An interesting 'LINEAR BLOCK ISARITHM EFFECT' of GPCP, created by a combination of the search method and the surface averaging technique, is illustrated in Figure 6.9. Where poorly distributed data occurs, particularly of a linear nature (eg. RIVER data), compressed linear isarithm blocks may occur (see bottom-right of centre in INCH-BREAKLINE) and may be aggravated even further by a 'LINEAR EFFECT' (see INCH-RIVER - centre-top and bottom-right-of-centre).

The 'LINEAR BLOCK ISARITHM EFFECT' occurs where all interpolated points on one side of the mid-point line (between two clusters of points - one high and one low) only use high points for their interpolation, whereas all points on the other side of the mid-point line only use low points (Figure 6.10a). This is a result of the search procedure selecting only the nearest 8 points, which, because the high and low

Figure 6.9



I-B GPCP

I-R GPCP

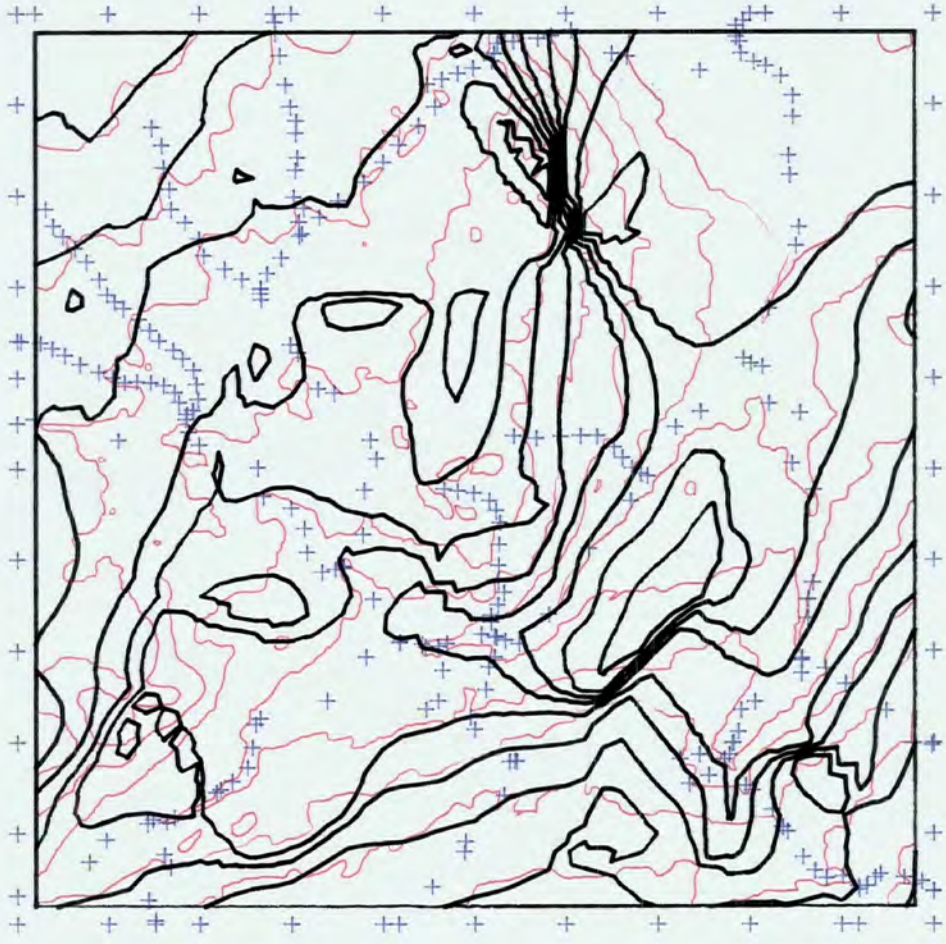
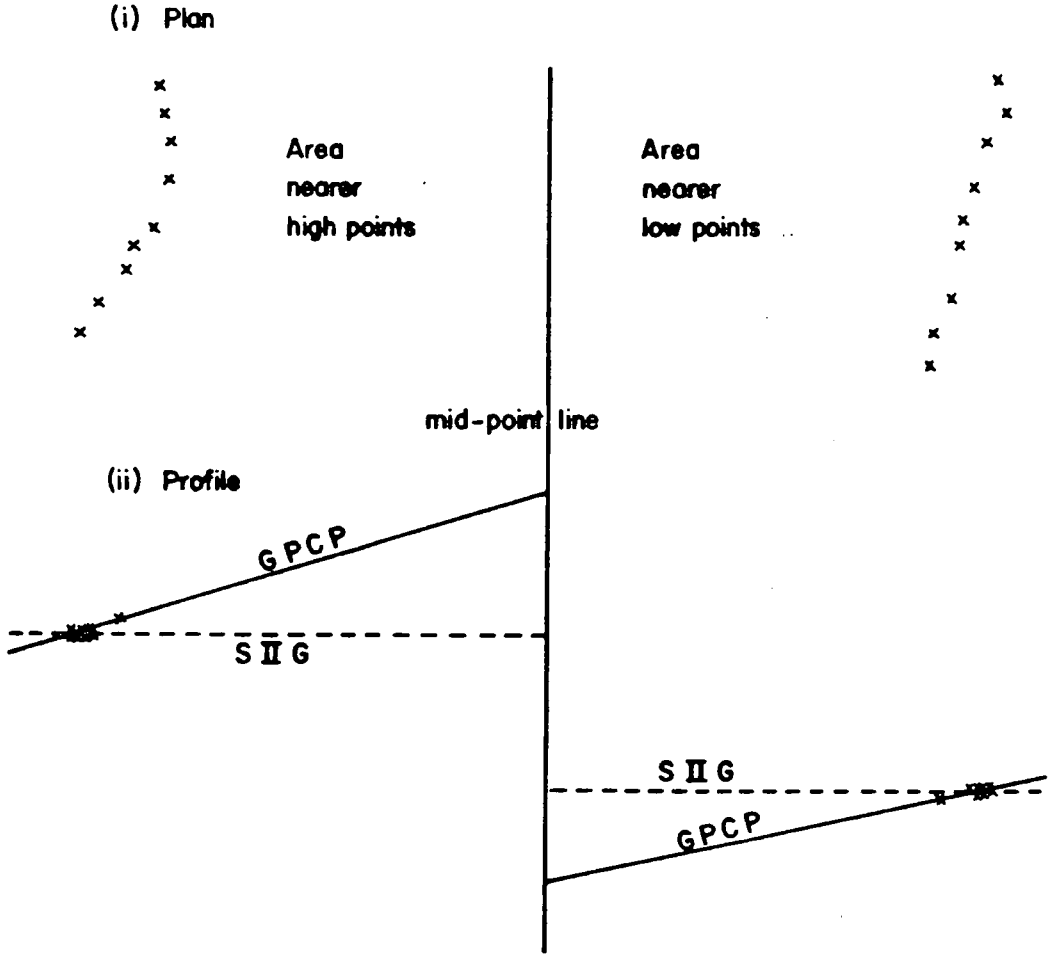


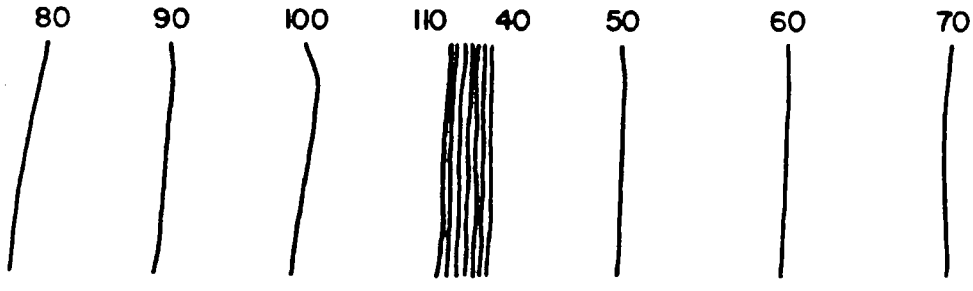
Figure 6.10

GPCP — LINEAR BLOCK ISARITHM EFFECT

a) Data points available



b) Resultant isarithms (GPCP)



points are in tight clusters, results in both clusters being mutually exclusive. This situation also occurs in SURFACE II GRAPHICS, but, as the interpolation function therein involves only a data averaging process, the resultant cliff is not so marked. GPCP further augments this effect when several points along a line are used to generate local surfaces at each of the points on the line ('LINEAR EFFECT'). Though a good surface approximation is made along the line, serious problems may occur across the line, and one rogue point may cause the local surfaces to dip or rise perpendicular to the line (see Figure 6.10.a.2). This linear effect therefore exaggerates the number of resultant isarithms (see Figure 6.10.b). Thus linear-type data with a low NNS (in fact linear data are generally not well described by NNS - a weakness of the technique - and comes out clustered), such as those from contours, may be considered totally unsuitable for a GPCP-type program which involves a 'two-stage' interpolation with a 'nearest n'-type search.

6.2.3 SYMAP

Within the SYMAP interpolation routine (see 2.5.5) the default search employed is similar to SURFACE II GRAPHICS-VRADIUS - the minimum radius being increased until between 4 and 14 points are found. Thereafter, the estimation process is similar to GPCP, but with the major exceptions that the weighting function includes a direction weight and that intermediate values derived in the two-stage process are restrained (as discussed in 2.5.5). As regards the weighting function,

> "the greater a data point's 'directional isolation' and the smaller its distance, the greater its weighting and hence its influence in determining the interpolated value" (Shepard, 1970, 5).

SYMAP generates an additional interpolation noise peculiar to this technique of computer mapping. This additional noise is a product of the data-to-pixel (cell) transformation inherent in a line-printer product.

> "Original sample points are assigned to grid cells or character locations. If more than one data point happens to fall inside a single grid cell, the value assigned to that area is the average of the data points" (Shepard, 1970, 7).

This averaged value is then used in the random-to-grid interpolation to interpolate grid nodes without involving the original data point. To overcome this problem, the original co-ordinates of the data points would have to be used rather than the cell value, which would increase the accuracy of the final product.

As previously, the consistency descriptors provide good indicators of the success of this method of interpolation (see Table 6.9). The %BAD statistics are very low - comparable with those for QUADRANT and VRADIUS which involved distant searching. This suggests that this might be a relatively safe method of interpolation. However, the ORIGINAL MIN and MAX statistics show that this method consistently produces a wide range of errors (mean magnitude - 22.11), much greater than even the largest equivalent SURFACE II GRAPHICS value (VRADIUS - 17.40). While this is marginally smaller than the equivalent value for GPCP (23.78), it must be considered poorer, since there were a few extreme magnitude values for GPCP which

DS	%BAD	ORIGINAL		TRIMMED		TRIMMED ERROR				TRIMMED ABSOLUTE ERROR			CORRELATIONS			
		MIN	MAX	MIN	MAX	MEAN	MEDIAN	SKEWNESS	KURTOSIS	MEAN	MEDIAN	STANDARD	SURFACE	NEAR	NONN	GRADIENT
IB	0.04	-34.78	40.86	-16.47	16.17	-0.59	-0.73	0.15	3.38	4.42	3.32	5.77	0.992	0.177	-0.116	0.247
IC	0.00	-21.69	33.27	-13.81	13.65	-1.69	-1.94	0.26	3.02	4.14	3.59	5.13	0.994	0.018	0.111	0.171
IG	0.00	-29.08	30.52	-12.85	12.81	-1.22	-1.39	0.24	2.93	4.15	3.50	5.22	0.993	0.079	-0.064	0.288
IR	1.96	-76.00	35.94	-30.61	30.13	-7.09	-6.68	0.15	3.93	9.21	7.77	11.63	0.980	0.386	-0.270	0.336
IS	0.04	-28.72	41.64	-14.68	14.64	-0.65	-0.47	0.05	3.31	4.02	3.10	5.23	0.994	0.121	-0.036	0.299
LB	0.00	-13.10	13.18	-4.75	4.68	-0.46	-0.39	-0.05	2.77	1.46	1.21	1.84	0.997	-0.019	-0.024	0.287
LC	0.00	-8.14	8.93	-4.31	4.31	-0.27	-0.44	0.26	2.57	1.50	1.30	1.84	0.996	0.147	-0.071	0.002
LG	0.00	-8.60	6.48	-3.59	3.42	-0.36	-0.28	-0.27	3.35	0.92	0.70	1.21	0.999	0.133	-0.046	0.372
LR	0.00	-22.86	25.13	-12.40	12.55	-0.84	-0.84	0.56	4.59	2.56	1.47	3.81	0.987	0.606	-0.473	0.520
LS	0.00	-12.43	11.18	-4.25	4.22	-0.44	-0.49	0.21	3.17	1.26	1.02	1.59	0.997	0.041	0.054	0.254
FB	0.00	-16.44	15.76	-9.69	9.75	0.76	0.35	-0.05	3.31	2.77	2.01	3.75	0.743	0.172	-0.122	0.281
FC	0.00	-15.63	13.35	-7.35	7.34	0.59	0.54	-0.31	3.17	2.22	1.77	2.88	0.831	-0.188	0.210	0.274
FG	0.00	-15.07	11.93	-7.31	7.31	0.48	0.41	-0.26	3.21	2.22	1.71	2.93	0.826	-0.026	0.087	0.328
FS	0.00	-13.60	14.64	-8.17	8.15	0.74	0.60	-0.22	3.18	2.44	1.93	3.18	0.804	0.072	-0.057	0.291

Table 6.9a Primary results of SYMAP

DS	IN-NEAR			IN-NONN			IN-GRAD			OUT-NEAR			OUT-NONN			OUT-GRAD		
	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN
IB	0.95	174.83	50.08	0	14	3.81	0.12	83.55	18.91	18.67	168.94	80.30	0	9	2.14	4.31	83.00	28.22
IC	0.83	182.47	44.10	0	9	3.29	0.12	83.55	18.85	7.39	157.25	50.05	0	10	4.22	1.00	81.55	31.75
IG	0.00	42.42	32.36	4	6	4.19	0.12	83.55	18.55	0.00	42.42	33.91	4	6	4.22	6.48	83.00	35.22
IR	0.91	306.41	70.75	0	20	3.46	0.12	83.55	18.72	53.97	297.06	164.24	0	8	0.25	3.81	78.75	31.91
IS	0.31	150.02	45.89	0	12	3.35	0.12	83.55	18.69	6.59	116.02	63.61	0	7	2.55	5.32	83.00	32.52
LB	0.31	165.00	39.92	0	23	3.17	0.00	41.50	10.06	3.63	85.39	41.21	0	20	2.64	4.14	30.24	16.75
LC	0.00	133.99	28.25	0	11	3.27	0.00	41.50	10.39	3.35	104.22	31.92	0	10	3.19	0.50	36.63	10.71
LG	0.00	42.42	19.53	1	6	3.34	0.00	40.00	10.01	0.00	42.42	25.13	1	5	3.05	2.23	41.50	17.83
LR	0.31	226.99	52.64	0	19	3.42	0.00	41.50	10.06	66.11	221.52	154.91	0	0	0.00	6.48	28.36	17.02
LS	0.21	112.92	27.41	0	11	3.05	0.00	41.50	10.01	3.26	74.33	32.83	0	9	2.92	3.21	36.63	17.99
FB	0.21	129.47	28.78	0	18	4.60	0.00	84.47	8.54	1.32	121.36	39.91	0	18	5.14	0.40	56.39	9.96
FC	0.40	70.00	21.72	0	15	4.46	0.00	84.47	8.35	1.13	39.24	15.30	1	17	8.14	1.07	56.60	13.47
FG	0.00	41.22	16.36	2	7	4.75	0.00	84.47	8.31	0.00	40.00	17.14	2	7	4.72	0.38	56.60	14.10
FS	0.00	84.33	24.83	0	17	4.60	0.00	84.47	8.38	0.00	48.52	20.83	1	12	5.60	0.90	56.39	12.89

Table 6.9b Secondary results of SYMAP

distorted the mean, whereas all the SYMAP MIN/MAX values have increased in magnitude at a fairly consistent rate. Another important aspect is that the MIN and MAX values are well-balanced and have a more similar magnitude than do any of those from the previous methods. This measure of bias will be discussed subsequently. In summary, all these aspects suggest that SYMAP is very much an averaging method of interpolation, which smooths the surface and inserts a large amount of random noise (perhaps a feature of the data-to-pixel noise discussed earlier).

The well-balanced ORIGINAL MIN/MAX statistics suggest that there is little interpolation bias. Yet, on a closer examination of the other bias descriptors, this is not seen to be the case. The MEAN and MEDIAN statistics are the most consistent within-model measures of bias encountered within the research. In addition, the MEAN and MEDIAN statistics for SYMAP mirror each other more closely than for the previous interpolation methods. This cannot be considered a random effect, as a very highly significant rank correlation exists between these statistics and mean altitude as derived within GEOD (MEAN - 0.95, MEDIAN - 0.96). The individual model systematic errors are more difficult to explain, but must be considered to be either features of the data sets, or characteristics of the relevant surfaces. Interestingly, the global mean MEANs for the two weighted surface averaging methods are quite similar (GPCP - -0.63, SYMAP - -0.79) and are distinct from the more direct weighted data average methods of SURFACE II GRAPHICS (approximately -0.15).

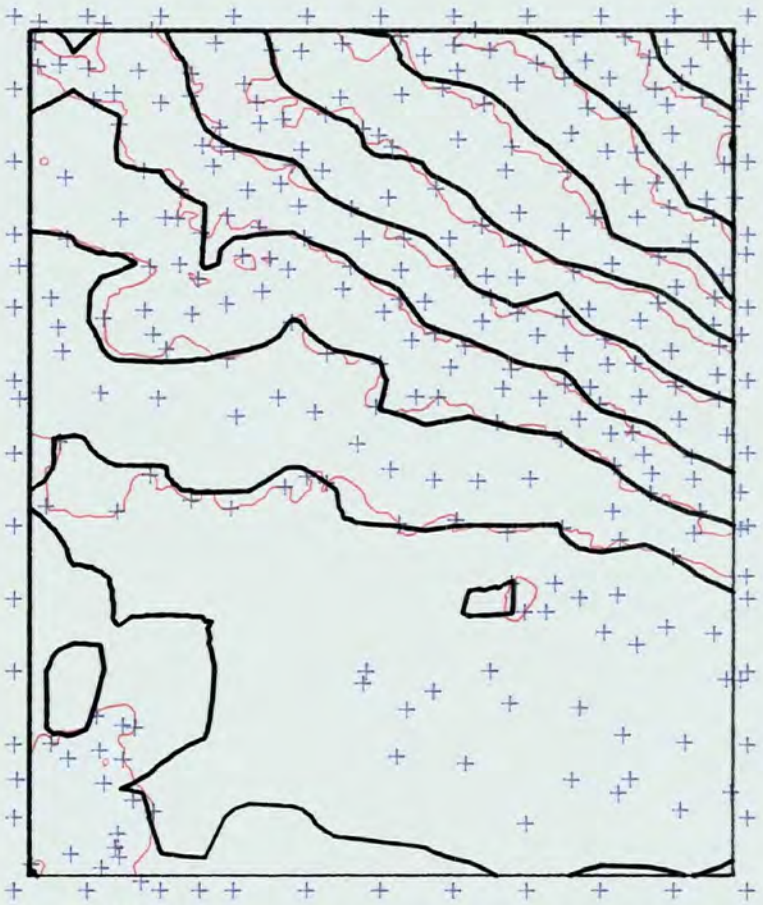
Greater consistency also occurs within the precision statistics. The TRIMMED MIN and MAX values show a marginally increased mean magnitude (GPCP - 9.73, SYMAP - 10.70) although there is much less within-model variance of the individual values, especially if the more awkward RIVER datasets are removed, which suffer from the problems of poor data distribution. Greater consistency is also very evident within the KURTOSIS values. A mean KURTOSIS of 3.28 is low compared with the previous interpolation methods and has low variance across the full range of data sets. This illustrates the near-normal distribution of error generated by SYMAP, especially if it is remembered that the TRIMMED limits of the error distribution are further apart, which would tend to marginally increaseKURTOSIS. This emphasises that the error generated by the SYMAP process must be considered more random than by any of the previous methods - possibly a direct result of the transformation of the data to the line printer format.

It is evident from Table 6.9 that the more sophisticated interpolation method of SYMAP promotes no improvement in resultant accuracy. Indeed, with the exclusion of the LIAN-BREAKLINE data, all STANDARD ERRORS, ABS MEAN errors and SURFACE CORRELATIONS are poorer than those for SURFACE II GRAPHICS and GPCP. The values for ABS MEAN and STANDARD ERROR indicate much greater within-model consistency, and this results in generally poor correlations with NNS (STANDARD ERROR - 0.80, ABS MEAN - 0.69 and SURFACE CORRELATION - 0.67). In addition, only weak secondary correlations exist - for example STANDARD ERROR with GEOD-derived gradient (0.55).

On examining the error-related correlations, very weak correlations are shown to exist for the NEAR (mean - 0.122) and NONN (mean - -0.058) statistics. With the exclusion of the very low NNS data sets (INCH-RIVER - 0.804 and LIAN-RIVER - 0.768), SYMAP may be considered to be an unbiassed data interpolator although, as demonstrated earlier, it is more suited towards terrain variability. This is supported by the ERROR/GRAD correlation coefficient which has a relatively high mean (0.268), with strong consistency. Additionally, the OUT-GRAD and IN-GRAD MEAN values show the importance of local gradient. All have good within-model consistency, and the OUT-GRAD MEAN values are distinctly greater than the IN-GRAD MEANs. Thus, by the removal of extremes caused by specific data point distribution characteristics, gross errors are removed, leaving only minor random errors (hidden by the general interpolation noise), which, in an averaging process, will be predominately based on gradient.

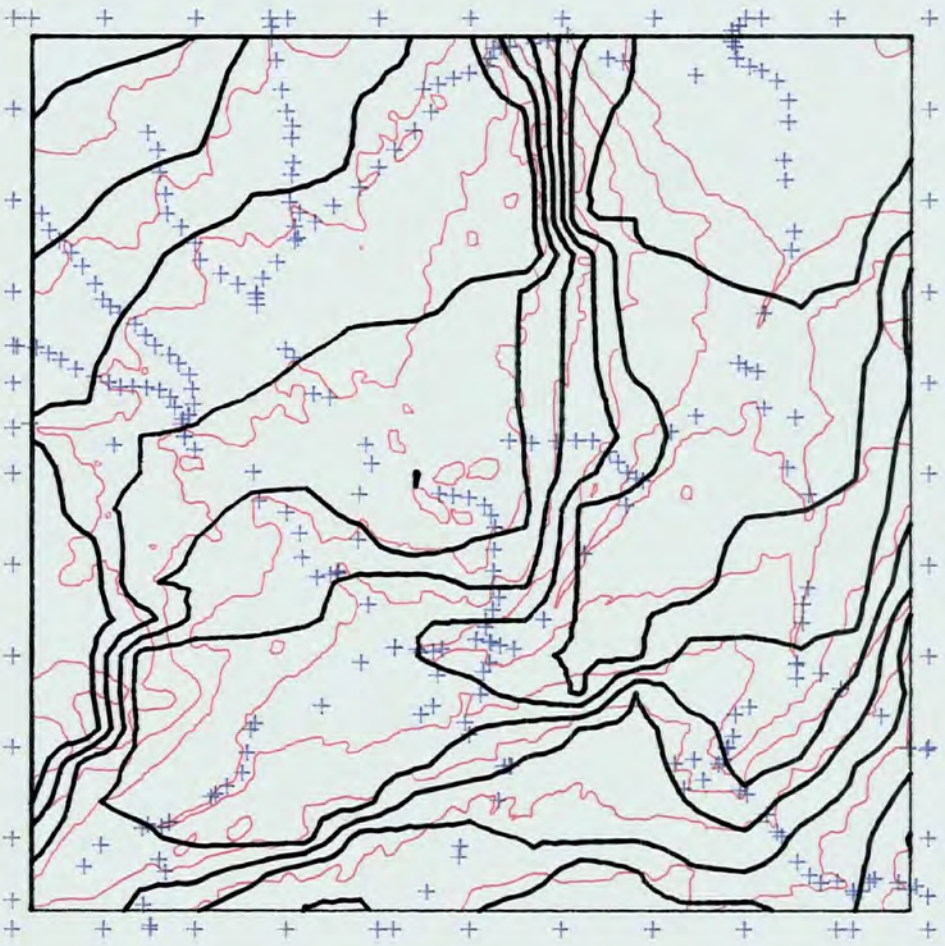
Although the summary statistics suggest that SYMAP is an inferior interpolator as compared with SURFACE II GRAPHICS and GPCP, a comparative examination of Figures 6.9 and 6.11 suggests that the morphological trueness is superior. Where a reasonable distribution occurs, a good correlation between interpolated and reference isarithms may be produced (LIAN-CONTOUR). Where a poor data distribution occurs, SYMAP has a superior ability to replicate the general trend of the model (eg. INCH-RIVER in Figures 6.9 and 6.11). The removal of extreme intermediate points has been stressed and thus the cliff (top-centre), while still in error, is not so extreme as with GPCP. Additionally, the edges of the cliff return much

Figure 6.11



L-C SYMAP

I-R SYMAP



more rapidly to reasonable values as a result of the directional weighting. At the cliff edges, perhaps seven high points and one low point are found, but since the seven high points are close together, they are each given a much smaller weight, thus increasing the importance of the lone, low value point. In this instance, the interpolated value will be lower and therefore closer to reality than, for example, with GPCP. This generally superior interpolation technique does, however, suffer from the data-to-pixel noise. Therefore, over the dense data areas, inferior quality isarithms may be generated unless a dense grid mesh is generated.

6.2.4 GINOSURF

The other common method of pointwise interpolation involves fitting a low-order polynomial (commonly quadratic) at each grid node. GINOSURF and GHOT utilise such methods, using distance-weighted least squares second-order polynomial fits. A major difference between the methods is that, whereas GHOT uses the full six term quadratic function, GINOSURF uses a four term paraboloid. Both methods use a 'nearest n' search, where n is an average value for the whole interpolation procedure. GINOSURF, which searches on average for the nearest 24 points, does not allow the number of points found to be less than 6 (always resulting in two redundancies), extending the search radius if necessary. GHOT searches on average for the nearest 30 points and does not extend the radius if too few points are found.

A striking aspect of the consistency descriptors is the

large and unbalanced nature of the ORIGINAL MIN and MAX values (see Table 6.10). In the case of LIAN-RIVER data, GINOSURF interpolates a maximum positive error of 99886.0 against a maximum negative error of -124.03. An interpolator which computes such a high error must be treated with a high degree of scepticism, and this is underlined by considering that there are seven ORIGINAL MIN and MAX values whose magnitude is greater than 50.0 for all the data sets. Indeed, excluding the LIAN-RIVER MAX value, the mean magnitude ORIGINAL error is 38.18 - over double the comparable SURFACE II GRAPHICS values and just under double the comparable values for the methods involving data surface averaging (GPCP and SYMAP). Similarly, the mean MIN (-49.82) and mean MAX (26.54) values are grossly unbalanced, suggesting a high negative bias. While there is a very highly significant (0.91) correlation between the individual mean magnitude ORIGINAL MIN/MAX values and NNS, this error cannot be blamed solely on NNS. There are several data sets which have extreme interpolation errors yet respectable NNS values (e.g. INCH-SCATTER - 33.23, LIAN-CONTOUR - -66.23 and FORV-SCATTER - -26.68). Notably, however, individual %BAD statistics adhere closely to previous values, suggesting that the extreme error is not indicative of all the interpolated values.

A sharp contrast is to be found in the precision statistics. The global mean magnitude TRIMMED MIN/MAX value (7.93) is only marginally bettered by OCTANT and QUADRANT. Indeed, if the LIAN-RIVER statistics were not so poor, the global mean could have been approximately 0.5 less than for any of the previous data sets. On closer examination, this

DS	%BAD	ORIGINAL		TRIMMED		TRIMMED ERROR				TRIMMED ABSOLUTE ERROR			CORRELATIONS			
		MIN	MAX	MIN	MAX	MEAN	MEDIAN	SKEWNESS	KURTOSIS	MEAN	MEDIAN	STANDARD	SURFACE	NEAR	NONN	GRADIENT
IB	0.08	-44.33	23.50	-10.88	11.02	0.50	0.58	-0.07	3.06	3.22	2.51	4.14	0.996	0.314	-0.227	0.092
IC	0.00	-27.44	18.11	-8.93	8.89	-0.26	-0.36	0.06	2.89	2.77	2.19	3.52	0.997	0.178	-0.012	0.001
IG	0.00	-14.10	16.88	-7.27	7.25	-0.10	-0.12	0.00	3.39	1.84	1.42	2.50	0.998	0.530	-0.451	0.141
IR	3.11	-139.63	62.02	-32.94	32.97	-4.97	-3.86	-0.16	5.15	7.22	5.00	10.00	0.983	0.417	-0.297	0.275
IS	0.00	-20.75	33.22	-10.30	10.47	0.90	0.76	-0.01	3.02	3.05	2.42	3.92	0.996	0.243	-0.050	0.131
LB	0.80	-162.80	85.58	-5.77	5.75	-0.65	-0.62	0.18	4.06	1.36	1.04	1.79	0.997	0.247	-0.232	0.010
LC	0.07	-66.22	29.08	-3.71	3.72	-0.10	-0.28	0.49	3.14	1.10	0.89	1.39	0.998	0.223	-0.199	-0.183
LG	0.00	-7.35	5.18	-2.22	2.22	-0.02	-0.10	-0.03	3.40	0.57	0.45	0.77	0.999	0.471	-0.085	0.140
LR	0.69	-124.02	99886.00	-10.71	10.81	-0.74	-0.71	0.38	4.90	2.31	1.43	3.31	0.990	0.559	-0.460	0.464
LS	0.00	-18.30	12.04	-3.14	3.13	-0.27	-0.32	0.23	3.01	0.96	0.80	1.21	0.998	0.156	-0.046	0.036
FB	0.00	-28.38	37.58	-5.80	6.06	0.08	0.08	0.67	5.57	1.05	0.70	1.52	0.959	0.236	-0.152	0.273
FC	0.00	-5.92	6.06	-2.44	2.44	0.12	0.08	0.09	3.15	0.68	0.54	0.88	0.986	0.100	0.114	0.108
FG	0.00	-11.55	6.34	-2.30	2.30	0.12	0.06	0.06	4.81	0.46	0.31	0.66	0.992	0.309	-0.056	0.251
FS	0.00	-26.67	9.38	-4.31	4.42	0.15	0.01	0.59	4.89	0.84	0.57	1.21	0.974	0.158	0.005	0.275

Table 6.10a Primary results of GINOSURF

DS	IN-NEAR			IN-NONN			IN-GRAD			OUT-NEAR			OUT-NONN			OUT-GRAD		
	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN
IB	0.95	171.50	49.85	0	14	3.84	0.12	83.55	19.22	9.01	174.83	84.89	0	8	1.60	0.89	83.00	22.50
IC	0.83	161.55	43.27	0	10	3.32	0.12	83.55	19.27	12.27	182.47	64.77	0	9	3.21	0.17	83.00	21.53
IG	0.00	42.42	32.16	4	6	4.21	0.17	83.00	18.96	30.00	42.42	37.66	4	5	4.02	0.12	83.55	27.38
IR	0.91	294.89	71.13	0	20	3.44	0.17	83.55	19.00	41.75	306.41	156.74	0	4	0.35	0.12	78.75	26.63
IS	0.31	150.02	45.67	0	12	3.35	0.12	83.55	18.99	21.83	148.00	67.80	0	7	2.60	2.90	77.96	26.92
LB	0.31	165.00	37.66	0	23	3.27	0.00	41.50	10.67	9.48	160.88	83.72	0	6	0.43	0.00	30.24	5.42
LC	0.00	133.99	27.52	0	11	3.35	0.00	41.50	10.61	7.11	125.03	45.61	0	8	1.68	0.00	27.72	6.27
LG	0.00	42.42	19.50	1	6	3.35	0.00	40.00	10.06	15.00	42.42	25.60	1	4	2.97	2.23	41.50	16.83
LR	0.31	226.99	52.99	0	19	3.40	0.00	41.50	10.02	18.88	226.99	147.85	0	3	0.08	6.82	35.99	17.55
LS	0.21	112.92	27.13	0	11	3.05	0.00	41.50	10.29	4.86	108.71	38.19	0	9	2.93	0.74	36.63	12.56
FB	0.21	129.47	27.66	0	18	4.81	0.00	84.47	8.60	12.85	129.05	61.08	0	10	1.40	0.13	61.88	8.72
FC	0.40	70.00	21.32	0	17	4.50	0.00	74.47	8.18	3.77	65.19	23.16	0	16	7.52	0.28	84.47	16.71
FG	0.00	41.22	16.22	2	7	4.77	0.00	84.47	7.86	10.00	40.00	20.00	2	7	4.43	1.07	74.47	22.60
FS	0.00	84.33	24.02	0	17	4.67	0.00	84.47	8.35	7.42	76.80	36.21	0	12	4.14	0.13	61.88	13.46

Table 6.10b Secondary results of GINOSURF

is seen to be a general trend amongst all data sets, with the exclusion of the very low NNS data sets.

This improvement in precision is strongly supported by the KURTOSIS statistics. The mean KURTOSIS (3.90) is higher than any previous mean. Its importance must be increased by the fact that SKEWNESS and the range of the error distribution have decreased, as compared with those from the other methods; both of these factors should lower KURTOSIS. These results, in conjunction with the previous consistency measures, suggest an interpolation method liable to produce occasional values in gross error, but which generally performs with a remarkably high precision.

The MEAN and MEDIAN statistics reflect few important features of interpolation bias. In general, all values have a greatly reduced magnitude to the previous interpolation methods, the mean systematic error is -0.38, lying between the weighted average methods (approximately -0.15) and the weighted data surface average methods (approximately -0.75). This suggests that the more symmetrical quadratic function has a greater stabilising effect than do the asymmetrical linear surface methods. A significant correlation (0.80) exists between MEAN error and the mean altitude (as computed by GEOD), although the similarity between the MEAN and MEDIAN values is more important since it suggests the TRIMMED error distribution is balanced and may be characterised by either measure of central tendency. This confirms the contrast between the few large unbalanced ORIGINAL extreme errors and the many precise, low error values.

The accuracy descriptors are similarly derived from the TRIMMED error distribution, and therefore the greatly increased accuracy statistics are not unexpected. The mean STANDARD ERROR (2.63) is markedly superior over the previous best value - QUADRANT (2.95). ABS MEAN and SURFACE CORRELATION have similarly improved, although, as the 0.96 (ABS MEAN and STANDARD ERROR/NNS) and 0.99 (SURFACE CORRELATION/NNS) correlations suggest, these are strongly related to NNS. In fact, the very low NNS RIVER data sets have marginally decreased in accuracy. This stresses that this is a generally superior method of interpolation within areas of good data distribution.

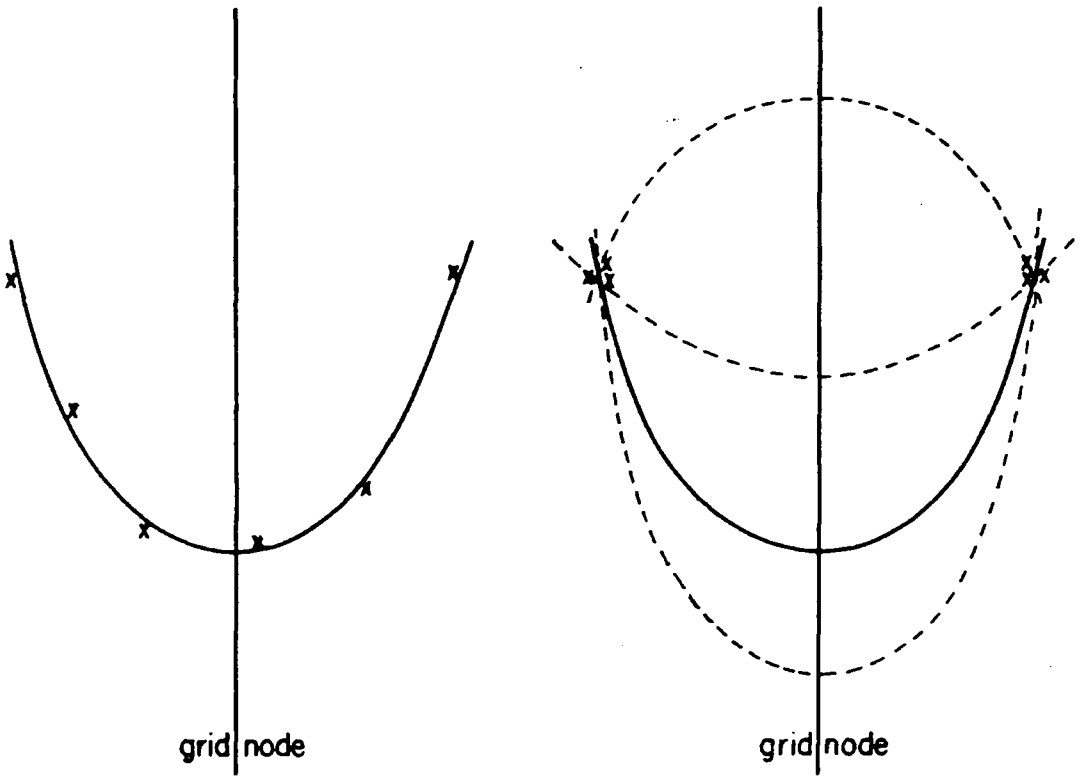
Interestingly, the NEAR error-related correlations which are weak overall display a bi-modal distribution with relation to NNS. The low NNS data sets will not be so successful in obtaining many redundant points in areas of sparse data for each point interpolation. Since more redundant points are likely to produce a better interpolation result, and a large range of data distributions and (therefore) redundancies will occur, a strong correlation might be expected. It is more difficult to suggest reasons for the high ERROR/NEAR correlations for high NNS data sets, although this is likely to be a result of polarisation of error. High NNS data sets produce low error by this method and therefore, where the NEAR value is 0.0, the error will be 0.0 (this frequently occurs in GRID data); elsewhere NEAR values rise to 42.43 (see OUT-NEAR for INCH-GRID and LIAN-GRID)increasing the error from 0.0 to modest values.

Overall, GINOSURF can perform with a remarkably high precision and accuracy, but may frequently, especially with poor distributions, generate extreme error values. These extremes are a result of the 'QUADRATIC EFFECT' and may be easily explained (see Figure 6.12) by considering the two-dimensional parabola. As points are more evenly distributed around the point of interest, so the parabola will be constrained (Figure 6.12a). However, where a poor data distribution occurs, with distant and with clustered data points around the grid node, the equation may become ill-conditioned - ambiguity being due to the data distribution. One point lower down the curve and nearer to the grid node can reduce this ambiguity. It is suggested that, from the data sets used within this research, data sets with nearest neighbour statistics of greater than 1.050 have generally a sufficiently even distibution to cope with this problem. The high kurtosis statistics are a product of the simplicity of the function and the number of neighbours used to interpolate a grid node. Assuming a few well-distributed points the fit should be good and thus the error should be low, generating a grid value close to reality. Since the average number of points to be found in the search is 24, the interpolating function should be well-constrained to the local environment, with the majority of interpolated values therefore having low errors and only a few having poorly interpolated extremes. This error distribution will produce high KURTOSIS statistics.

All these points are emphasised in Figures 6.13 and 6.14. It should be noted that, in these and the GHOT maps, isarithms

Figure 6.12

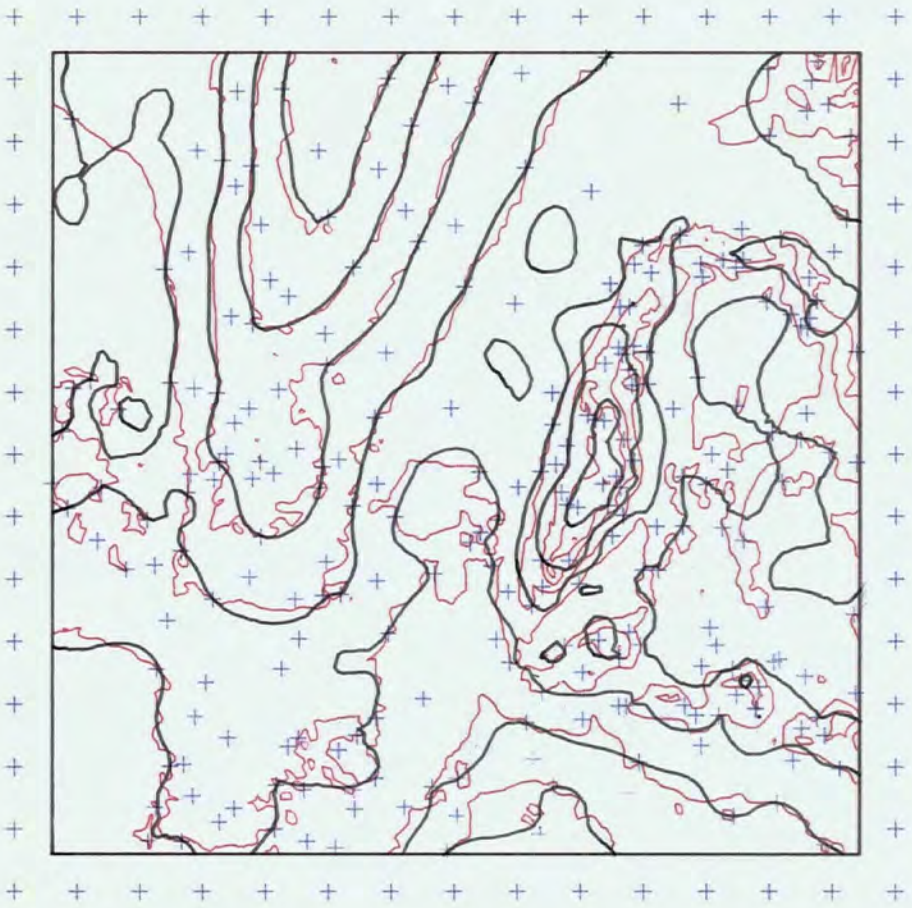
The quadratic effect



a) constrained by a good distribution.

b) ill-conditioning due to poor distribution results in many possible curves.

Figure 6.13



F-C GINO

F-B GINO

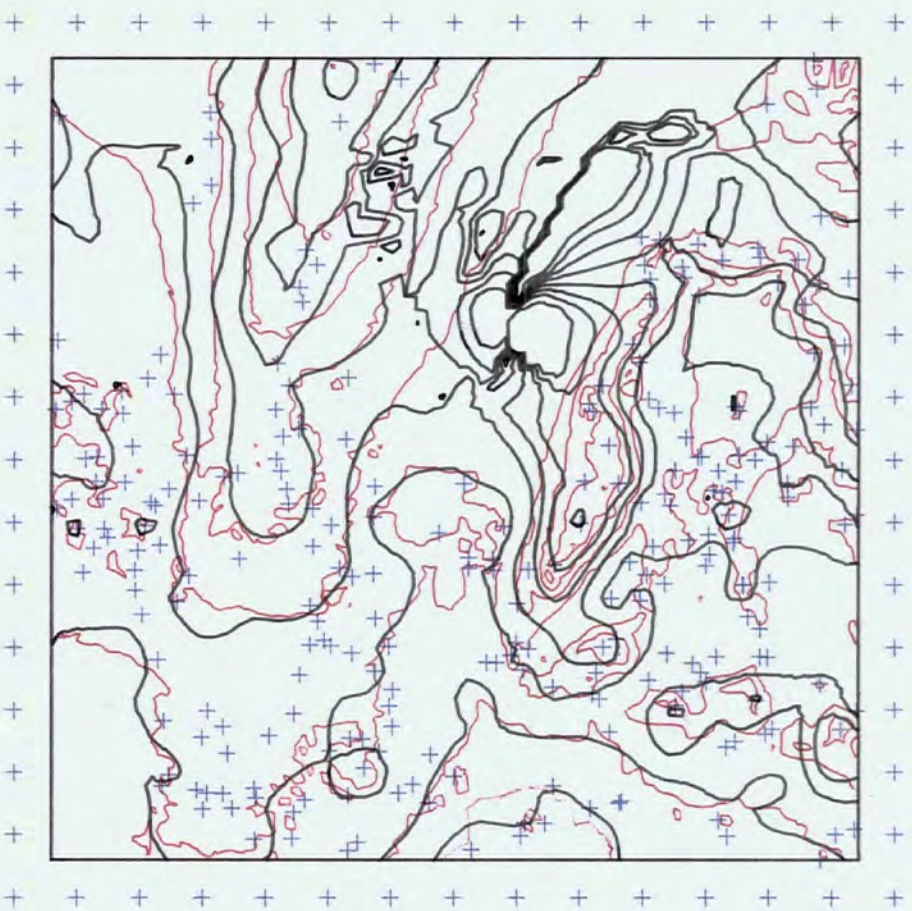
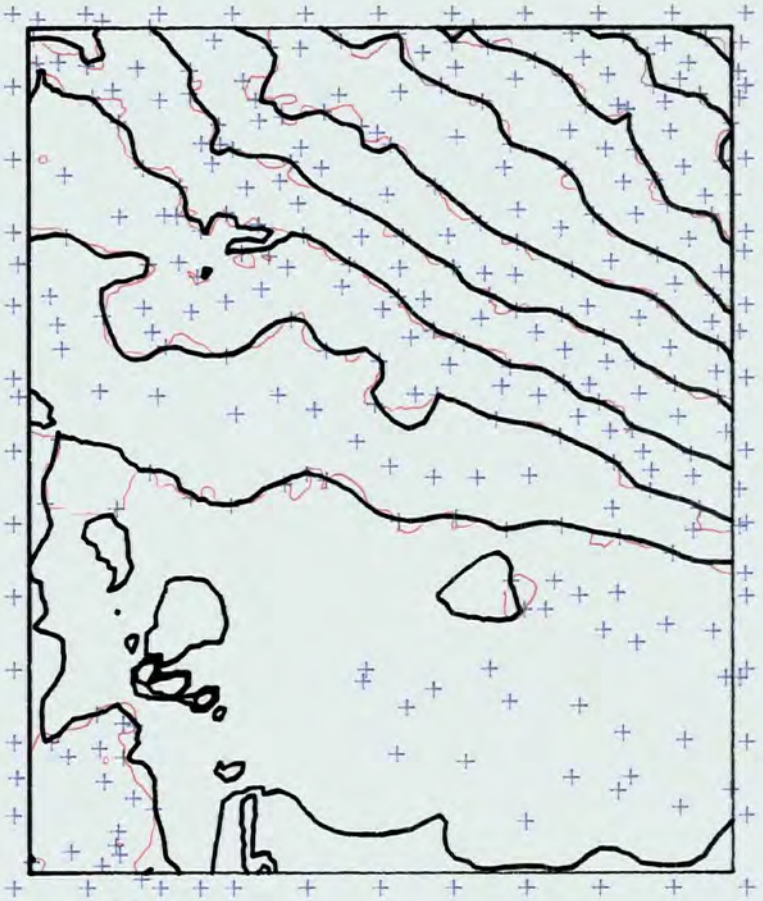
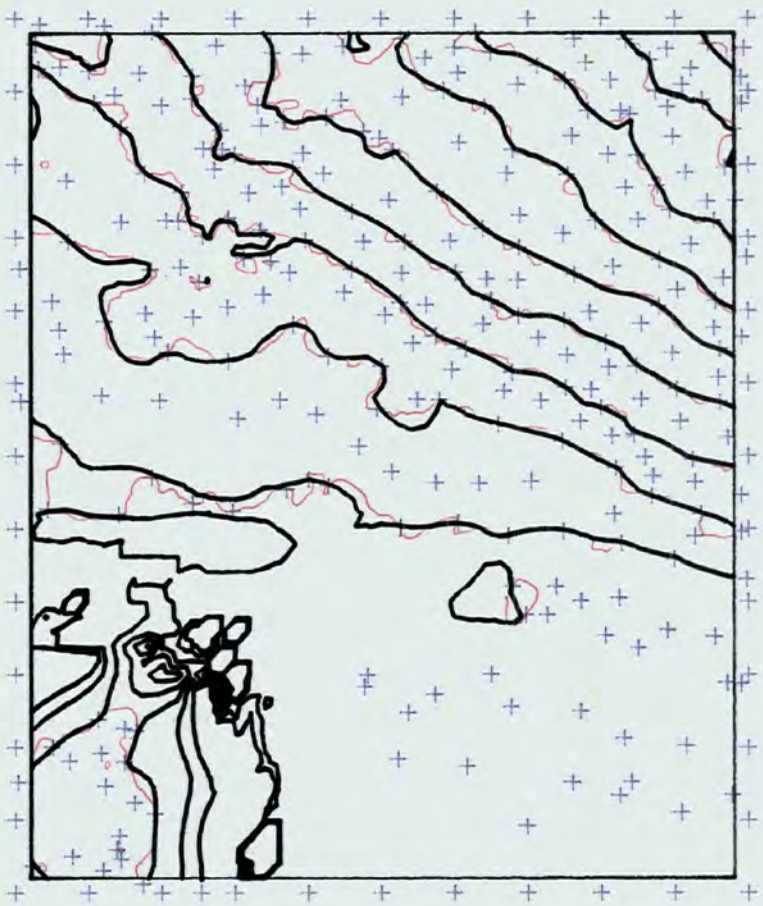


Figure 6.14



L-C GINO

L-C GHOT



were not plotted at levels outside a reasonable range of the reference isarithms (for reasons discussed in 6.1.2). In general, the resultant isarithms are superior to all methods of interpolation previously discussed. Compare, for example, the superior FORV-GRID distribution in Figures 6.3 and 6.5 with the poorer FORV-CONTOUR distribution in Figure 6.13, and similarly compare the alternative FORV-BREAKLINE (Figures 6.13 and Figure 6.6) results. There is no doubt that using LIAN-CONTOUR data, GINOSURF (Figures 6.14) produces closer fitting isarithms than does SYMAP (Figure 6.11). It must therefore be assumed that a second-order polynomial is a modelling tool more suited to the data sets examined. However, serious problems occur in data-voids (see Figure 6.13b) and thus ridiculous results emerge in such areas with extreme maxima and minima (e.g. LIAN-RIVER in Table 6.10).

6.2.5 GHOT

A comparison of the previous GINOSURF summary statistics (Table 6.10) and the GHOT statistics (Table 6.11) suggests that the increased number of polynomial coefficients used by GHOT (2) and its poorer search routine have a dramatic effect. The %BAD statistics have increased considerably though, on close examination, it was noted that approximately 75% of the nodes badly estimated by GINOSURF were satisfactorily estimated by GHOT. Interestingly, the %BAD statistic has only a significant (0.81) correlation with NNS and areal autocorrelation (AUFN). This is one of the lowest NNS

DS	%BAD	ORIGINAL		TRIMMED		TRIMMED ERROR				TRIMMED ABSOLUTE ERROR			CORRELATIONS			
		MIN	MAX	MIN	MAX	MEAN	MEDIAN	SKEWNESS	KURTOSIS	MEAN	MEDIAN	STANDARD	SURFACE	NEAR	NONN	GRADIENT
IB	0.42	-36.61	389646.00	-11.39	11.43	0.39	0.46	-0.08	3.13	3.26	2.55	4.19	0.995	0.276	-0.202	0.085
IC	0.33	-37.42	42.89	-26.05	26.00	-1.01	-0.78	-0.01	2.18	9.97	9.31	11.93	0.965	0.090	-0.143	-0.008
IG	0.00	-14.09	16.83	-7.26	7.19	-0.10	-0.42	-0.03	3.39	1.82	1.40	2.48	0.998	0.529	-0.453	0.138
IR	7.54	-229.63	36937.05	-48.35	47.44	-7.61	-5.22	-0.89	6.22	9.18	5.89	13.48	0.973	0.480	-0.397	0.291
IS	0.00	-19.46	30.69	-9.92	9.84	0.83	0.76	-0.10	3.01	2.93	2.39	3.73	0.997	0.211	-0.029	0.124
LB	4.64	-160.83	20669.00	-33.46	29.69	-2.00	-1.47	-2.07	13.18	3.14	1.52	6.13	0.971	0.392	-0.380	-0.282
LC	1.25	-1040.70	651.80	-6.94	6.15	-0.43	-0.47	0.01	5.18	1.19	0.92	1.61	0.997	0.285	-0.275	-0.230
LG	0.00	-7.22	5.27	-2.84	2.75	-0.02	-0.09	-0.08	4.02	0.61	0.45	0.84	0.999	0.451	-0.085	0.174
LR	4.64	-179.16	1656.90	-34.19	34.55	-2.10	-1.56	0.44	8.65	4.65	2.10	7.71	0.948	0.516	-0.547	0.362
LS	0.10	-113.28	33.75	-3.30	3.31	-0.27	-0.30	0.22	3.14	0.95	0.80	1.18	0.998	0.147	-0.045	0.045
FB	1.92	-129.05	311.50	-12.15	12.21	0.38	0.00	0.76	7.94	1.61	0.91	2.63	0.898	0.417	-0.322	0.167
FC	0.00	-6.39	6.59	-2.38	2.27	0.14	0.11	0.09	3.06	0.69	0.53	0.89	0.987	0.110	0.107	0.102
FG	0.00	-11.25	5.97	-2.30	2.32	0.13	0.05	0.06	4.76	0.44	0.30	0.65	0.992	0.306	-0.054	0.255
FS	0.00	-38.78	33.52	-14.13	9.59	0.23	0.07	-0.53	11.43	1.21	0.68	2.02	0.930	0.239	-0.069	0.192

Table 6.11a Primary results of GHOT

DS	IN-NEAR			IN-NONN			IN-GRAD			OUT-NEAR			OUT-NONN			OUT-GRAD		
	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN
IB	0.95	174.83	50.16	0	14	3.84	0.12	83.55	19.42	9.01	162.11	77.27	0	8	1.76	0.64	78.75	18.61
IC	0.83	182.47	43.97	0	9	3.31	0.12	83.55	19.13	4.00	143.71	51.42	0	10	3.57	1.68	65.32	24.05
IG	0.00	42.42	32.16	4	6	4.21	0.17	83.00	18.97	30.00	42.42	37.46	4	5	4.02	0.12	83.55	27.13
IR	0.91	306.41	70.66	0	20	3.47	0.12	83.55	19.02	30.00	245.80	157.30	0	2	0.10	2.43	68.89	25.50
IS	0.31	150.02	45.86	0	12	3.35	0.12	83.55	19.00	17.05	130.97	64.36	0	7	2.80	2.90	77.96	26.69
LB	0.31	165.00	38.28	0	23	3.29	0.00	41.50	10.69	15.00	136.85	70.41	0	2	0.50	0.00	27.05	5.10
LC	0.00	133.99	26.72	0	11	3.39	0.00	41.50	10.72	15.00	125.03	60.96	0	4	0.72	0.26	16.80	4.27
LG	0.00	42.42	19.63	1	6	3.34	0.00	40.00	10.21	15.00	42.42	26.47	1	4	2.84	2.69	41.50	17.16
LR	0.31	226.99	55.74	0	19	3.40	0.00	41.50	10.30	15.00	182.47	94.17	0	2	0.32	1.76	31.02	12.27
LS	0.21	110.58	26.97	0	11	3.06	0.00	41.50	10.36	4.86	112.92	41.44	0	9	2.61	0.00	36.63	11.06
FB	0.21	129.47	27.50	0	18	4.85	0.00	84.47	8.77	29.75	112.10	64.39	0	3	0.49	0.56	17.94	5.35
FC	0.40	70.00	21.55	0	17	4.44	0.00	74.47	8.10	3.77	59.67	18.77	0	16	8.52	0.45	84.47	18.25
FG	0.00	41.22	16.22	2	7	4.77	0.00	84.47	7.89	10.00	40.00	20.05	2	7	4.42	1.07	74.47	22.33
FS	0.00	84.33	24.61	0	17	4.64	0.00	84.47	8.61	30.00	41.22	37.05	1	3	1.56	1.27	3.88	2.64

Table 6.11b Secondary results of GHOT

correlations and one of the highest secondary correlations observed, underlining the dramatic nature of this interpolation method.

More seriously, there are twelve excessively extreme ORIGINAL MIN/MAX values of which three (INCH-BREAKLINE - 389646., INCH-RIVER - 36937. and LIAN-BREAKLINE - 20669.) must be considered ridiculous. Even if these three values are not considered, the mean magnitude ORIGINAL MIN/MAX value is 93.22 - approximately 3x greater than the value for GHOT and approximately 5x greater than each of the preceding values. However, there are five data sets which have marginally improved (all GRID data sets and INCH-SCATTER and FORV-CONTOUR). While these include the data sets with the highest NNS in each model, they also include data sets with high autocorrelation, stressing the dual contributors to this method of interpolation. Additionally, this technique displays a marked positive bias to the ORIGINAL positive and negative extremes.

Unlike GINOSURF, the areas of extreme error are not small and distinct from the large areas of precise interpolation. Precision has been similarly affected, with a global TRIMMED MIN/MAX mean magnitude of 14.93 (as against approximately 8.00 to 9.00 for most other methods). In GINOSURF, the extreme errors were confined to the 0-2.5% and 97.5-100% quantile ranges of the error distribution. However, in GHOT, these large errors extend into the 2.5-97.5% error distribution range, stressing the unreliable nature of the interpolator.

This encroachment of extreme errors into the TRIMMED

MIN/MAX range has without doubt affected KURTOSIS, although increased SKEWNESS is partly responsible for the increased LIAN-BREAKLINE KURTOSIS. The relatively large TRIMMED MIN/MAX ranges produced from the INCH-RIVER, LIAN-BREAKLINE, LIAN-RIVER, FORV-BREAKLINE and FORV-SCATTER data sets are matched by large KURTOSIS values (6.22, 13.19, 8.66, 7.95 and 11.43). These clearly have contributed to the large mean KURTOSIS (5.67), although it should be noted that virtually all individual values have increased on the (previous high) GINOSURF values and, for example, LIAN-GRID (4.02) and LIAN-CONTOUR (5.18) data sets do not give rise to either large SKEWNESS or large TRIMMED MIN/MAX ranges. It must therefore be concluded that this method has increased the previous large errors while being a more precise interpolator in the low error situations.

Systematic error has behaved similarly to that from GINOSURF, although it has produced stronger values. The global mean MEAN has increased in magnitude from -0.38 to -0.82, although this can be mostly accounted for in three data sets - INCH-RIVER, LIAN-BREAKLINE and LIAN-RIVER. Interestingly, there is a higher MEAN/altitude correlation (0.93) than for GINOSURF (0.80). This is an inherent feature of using a quadratic interpolation function within a pointwise method - since it ensures that lows are lower and highs are higher. By examining Figure 6.12, it is obvious that the curved part of the parabola (which occurs at the point to be interpolated) must either dip below or rise above the surrounding data. In a situation such as that contained in RIVER data, where the data are already biassed low, the resultant interpolated grid will

be more strongly biassed low - and vice versa.

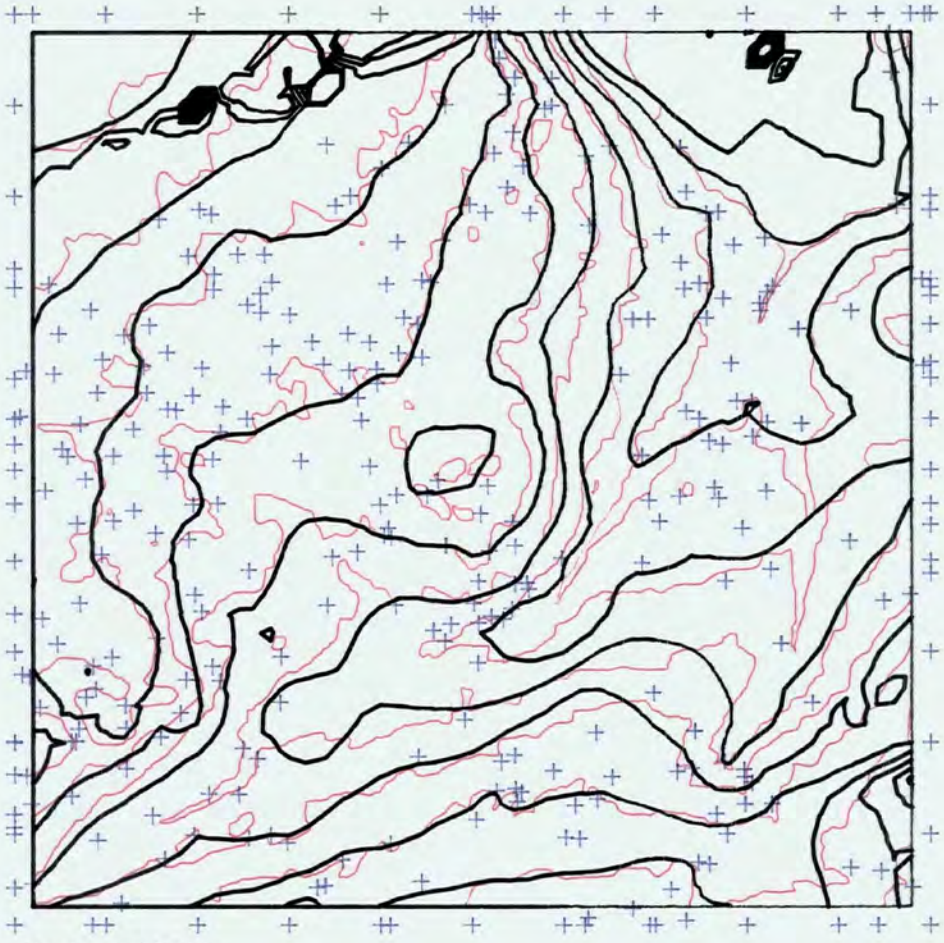
The contrast between the sharp decrease in the strength of the interpolator for some data sets, and the similar, sometimes minor increases gained by using this method over GINOSURF, are illustrated in the accuracy descriptors. Use of the INCH-CONTOUR, INCH-RIVER, LIAN-BREAKLINE, LIAN-RIVER and FORV-BREAKLINE data sets shows sharp decreases in accuracy. Alternatively, use of the INCH-GRID, INCH-SCATTER, LIAN-SCATTER and FORV-GRID data sets shows minor increases in accuracy. It is much more difficult to suggest the precise reasons which give rise to high or low accuracies for different data sets. Accuracy/NNS correlations are much lower than for previous methods (ABS MEAN, and SURFACE CORRELATION - 0.75 and STANDARD ERROR - 0.86). However, there has been a dramatic rise in accuracy/areal autocorrelation correlations (ABS MEAN, SURFACE CORRELATION are very highly significant - 0.94 and STANDARD ERROR is highly significant 0.90). On a non-statistical basis, it is apparent that the lineated, surface-specific data sets (BREAKLINE, RIVER and CONTOUR) have fared dramatically less well than the surface-random data sets (GRID and SCATTER).

This conflict is apparent within the error-related correlations and IN and OUT summary statistics. While mean values fluctuate apparently importantly (ERROR/NEAR - 0.317, ERROR/NONN - -0.207 and ERROR/GRAD - 0.101), the within-model variances are very large.

The view that surface autocorrelation is always of major significance to the resultant accuracy of interpolation may be

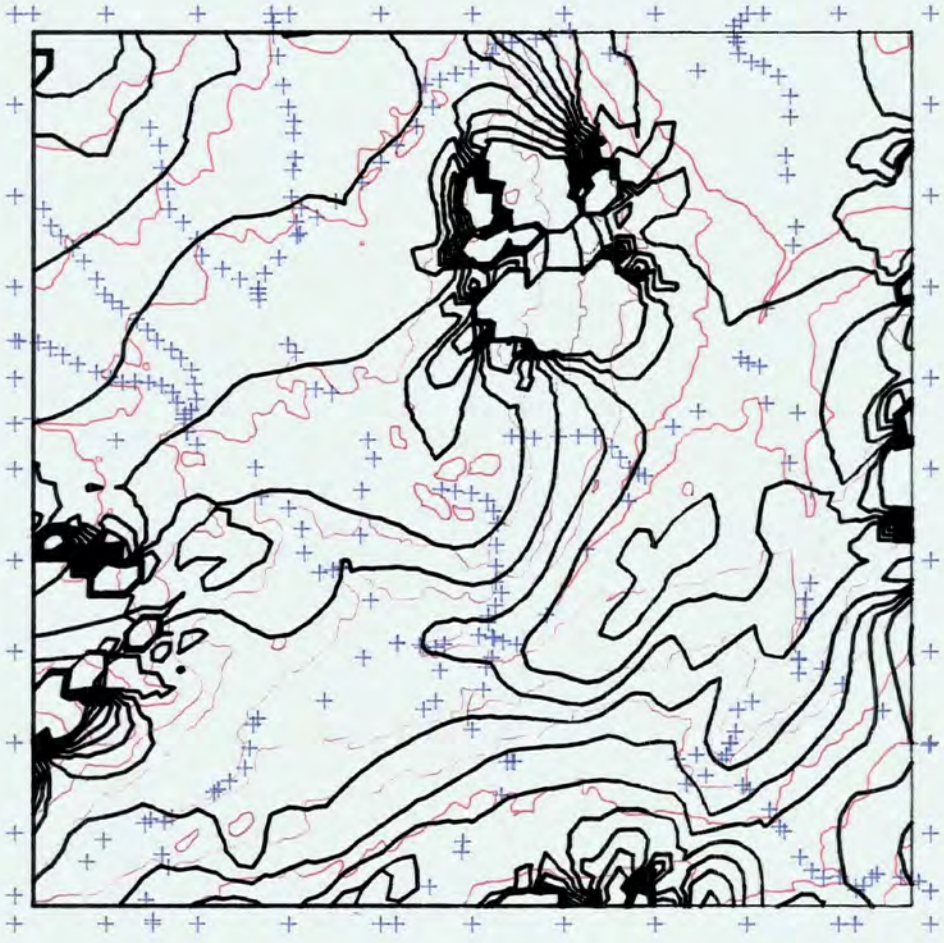
weakened by considering Figure 6.14 with respect to Figure 6.11. There is no doubt that, outside the flat river valley in the lower third of the area, the isarithms are well-replicated both by GHOT and GINOSURF. However, in the flat, highly autocorrelated area (well interpolated by SYMAP), extremes of -1040m and +650m result in very poor resultant isarithms for GHOT, distinctly inferior even to those from GINOSURF. This must be considered a direct consequence of a low value for the NNS and irrespective of any other autocorrelation or surface area or dispersion vector statistics. This is stressed by comparing Figure 6,9 (the GPCP comparison) with Figure 6.15. Outside data voids, GHOT produces superior results to any other pointwise methods. However, a program is only of practical use if it always produces reasonable results, or suppresses output where it is unable to achieve reasonable results. Extremes of 389,646m and 36,937m- factors of 2000x and 200x the range of the data - are excessive for any interpolation method and render it unreliable, especially since the quadratic effect and ill-conditioning do not restrict the error to one small location, but extend it to large blocks of grid nodes. The general marginal increase in error/number of near neighbours does demonstrate the importance of the ill-conditioning due to insufficient points. When any polynomial function is utilised, the number of points which must be found is equivalent to the number of coefficients. In a least squares fit, this number may be augmented. There is no guarantee that, within GHOT, the minimum number of points is used for each evaluation, hence the slight increase in error/number of near neighbours

Figure 6.15



I-B GHOT

I-R GHOT



correlation since fewer points result in a poorer fit. Additionally, since the six-term quadratic function utilised in GHOT is more flexible and therefore potentially more unstable than the four-term paraboloid used in GINOSURF, a poor point distribution may incur a larger subsequent error in GHOT.

Overall, the second-order function can, in general, model the local nuances of terrain better than a weighted average or linear surface provided a reasonable point distribution exists. This is emphasised by the error/gradient correlations being found to be trivial for both GHOT and GINOSURF and much weaker than the error/nearest neighbour distance. GHOT may be summarised as performing with reasonable accuracy, but producing many extreme values. The problem of data distribution appears to be more crucial than with GINOSURF and a marked correlation exists between high accuracy and high NNS. Thus, it is suggested that, from the results of the data sets used in this research, only data sets with a NNS greater than 1.350 can be guaranteed to produce sufficiently consistent results. The main sources of error in GHOT are the poor search procedure and the less stable function.

6.2.6 MINCURV

Within the two-stage process of interpolation using MINCURV (see 2.5.8) there are several parameters which must be defined. The search radius (s) used in the weighted average interpolation to establish the initial grid was generously defined to prevent any gross preliminary grid errors. The

minimum successive iteration difference (DELTA) at which iterations should terminate was based on the original digitising height accuracy and defined as 0.25m. The starting grid dimensions obviously had to vary to suit the particular model concerned, since the final grid dimensions would differ. They were defined to allow four successive grid expansions for each model before the final required size was derived. The other necessary parameters were suitably defined.

In the program description of MINCURV (Swain, 1976), it is stated that,

> "Results are considered superior to those of standard packages such as Calcomp's GPCP because the surface produced has the minimum curvature property."

However, from the results of the previous packages (including GPCP), it is obvious that, within the considerations of this research, the minimum curvature method is distinctly inferior to all previously discussed methods (see Table 6.12). MINCURV has a very poor estimation reliability; an average of 14.8% of grid nodes within INCH and LIAN data sets were not evaluated within the required tolerances. It is unlikely that the poor interpolation is due to poor parameter selection, since the majority of the parameters are associated with the establishment of the preliminary grid which should not create gross errors. Within each model, interpolation accuracy has a very highly significant correlation (0.96) with NNS, although importantly, all grid nodes within FORV data sets are satisfactorily evaluated. This may be related to the smoother terrain of FORV, as defined by the vector analysis and surface area parameters, although probably is also due to the consistently

DS	%BAD	ORIGINAL		TRIMMED		TRIMMED ERROR				TRIMMED ABSOLUTE ERROR			CORRELATIONS			
		MIN	MAX	MIN	MAX	MEAN	MEDIAN	SKEWNESS	KURTOSIS	MEAN	MEDIAN	STANDARD	SURFACE	NEAR	NONN	GRADIENT
IB	14.57	-278.52	277.04	-126.07	124.80	-4.01	-0.99	-0.78	9.97	13.54	4.69	25.55	0.871	0.547	-0.528	-0.154
IC	5.57	-281.60	161.08	-45.86	45.88	-1.02	-0.44	-1.21	12.25	4.76	2.81	8.09	0.983	0.307	-0.235	-0.196
IG	0.17	-28.13	57.38	-27.66	27.55	-0.27	-0.63	-0.27	10.59	2.71	1.71	4.34	0.995	0.405	-0.357	0.100
IR	36.74	-342.28	517.56	-194.63	193.44	-7.14	-7.27	0.43	4.93	37.11	16.08	57.52	0.660	0.734	-0.706	0.112
IS	3.75	-140.17	73.86	-35.03	34.78	-0.61	0.57	-0.92	6.61	5.57	3.44	8.47	0.980	0.430	-0.330	0.013
LB	31.13	-118.80	216.41	-118.80	190.14	21.33	-5.75	1.39	4.09	35.58	9.71	62.21	-0.096	0.759	-0.695	-0.547
LC	17.22	-22.07	339.82	-22.07	189.88	14.56	-0.96	2.76	9.80	16.47	2.04	41.85	0.233	0.688	-0.695	-0.501
LG	1.06	-41.64	134.66	-16.60	16.22	0.00	-0.14	-0.31	12.56	1.63	0.69	3.23	0.989	0.366	-0.239	-0.096
LR	46.77	-128.72	299.87	-128.72	288.11	-2.01	-20.92	1.34	4.71	51.86	32.83	77.30	-0.293	0.593	-0.725	0.149
LS	8.02	-63.36	149.49	-63.36	86.77	2.22	-1.19	2.43	13.11	6.97	1.93	15.84	0.762	0.444	-0.489	-0.330
FB	0.00	-5.29	10.93	-4.47	4.71	0.54	0.29	0.64	3.42	1.14	0.78	1.57	0.962	0.395	-0.250	0.183
FC	0.00	-6.60	5.89	-2.38	2.39	0.30	0.28	-0.06	2.89	0.73	0.61	0.92	0.986	0.158	-0.062	0.067
FG	0.00	-10.64	8.52	-2.38	2.39	0.13	0.08	0.02	4.77	0.49	0.34	0.70	0.991	0.329	-0.103	0.243
FS	0.00	-8.02	7.02	-3.81	3.88	0.32	0.16	0.51	3.96	0.85	0.56	1.18	0.975	0.153	0.019	0.146

Table 612a Primary results of MINCURV

DS	IN-NEAR			IN-NONN			IN-GRAD			OUT-NEAR			OUT-NONN			OUT-GRAD		
	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN
IB	0.95	174.83	49.78	0	14	3.86	0.12	83.55	19.72	4.38	25.35	12.75	0	5	1.02	20.33	68.94	85.71
IC	0.83	158.83	42.19	0	10	3.43	0.12	83.55	20.05	29.66	182.47	85.05	0	4	1.22	0.56	34.17	6.52
IG	0.00	42.42	32.41	4	6	4.19	0.12	83.55	19.36	30.00	42.42	38.27	4	4	4.00	13.64	32.88	24.91
IR	0.91	306.41	74.89	0	20	3.42	0.12	83.55	19.36	30.00	189.74	85.25	0	3	0.93	2.64	52.33	19.60
IS	0.31	144.32	45.66	0	12	3.40	0.12	83.55	19.75	18.02	150.02	68.13	0	5	1.69	0.56	33.50	12.14
LB	0.31	160.36	36.00	0	23	3.30	0.00	41.50	10.75	33.53	165.00	115.97	0	2	0.03	0.56	9.43	3.73
LC	0.00	125.03	26.13	0	11	3.40	0.00	41.50	10.72	15.00	133.99	72.38	0	2	0.42	0.50	16.80	4.42
LG	0.00	42.42	19.47	1	6	3.35	0.00	40.00	10.50	15.00	42.42	26.47	2	4	2.80	0.00	41.50	8.59
LS	0.31	226.99	57.77	0	19	3.25	0.00	41.50	10.39	21.21	60.00	38.25	0	2	1.00	4.68	16.47	9.46
LS	0.21	86.11	26.07	0	11	3.14	0.00	41.50	10.60	21.21	112.92	77.60	0	2	0.04	0.56	9.35	3.93
FB	0.21	129.47	27.94	0	18	4.77	0.00	74.47	8.51	9.97	119.94	55.88	0	13	2.14	0.13	84.47	10.43
FC	0.40	70.00	21.44	0	16	4.47	0.00	84.47	8.21	2.47	44.72	20.69	1	17	7.81	0.45	61.88	16.14
FG	0.00	41.22	16.19	2	7	4.77	0.00	84.47	7.96	10.00	40.00	20.67	2	7	4.35	1.07	74.47	21.02
FS	0.00	84.33	24.55	0	17	4.59	0.00	84.47	8.17	6.54	58.88	26.24	0	13	5.60	0.90	61.88	16.86

Table 6.12b Secondary results of MINCURV

superior data distributions within the FORV model.

All error statistics are also consistently proportionately larger for INCH and LIAN. Maximum and minimum errors are extreme and often unbalanced, although the resultant systematic error has a lower magnitude, since it does not include the extreme values. The unpredictability is emphasised by KURTOSIS, where within each model there is a large variance (approximately 4-14). The importance of the data distribution is further stressed by the error correlation statistics. In no other interpolation method examined do such consistent, very highly significant ERROR/NEAR (mean - 0.451) and ERROR/NONN (mean - -0.385) correlations exist. The lack of any influence of the surface on the interpolation procedure is suggested by the comparatively poor ERROR/GRAD correlations which have a large within-model variance.

While it is trivial to point out the overall cause of error, namely poor data distribution, it is much more complex to explain the precise reason for its occurrence due to the interwoven nature of the interpolation procedure. Swain (1976, 231) discusses some of the problems involved in determining the initial grid which is subsequently utilised in the minimum curvature iteration stage. This initial random-to-grid stage may be performed in several ways, although,

> "Slightly faster convergence of the solution may be obtained by initially fitting weighted paraboloids at each grid point, although this can give extreme values in large gaps and at the edges where all the data are on one side of a grid point" (Swain, 1976, 231).

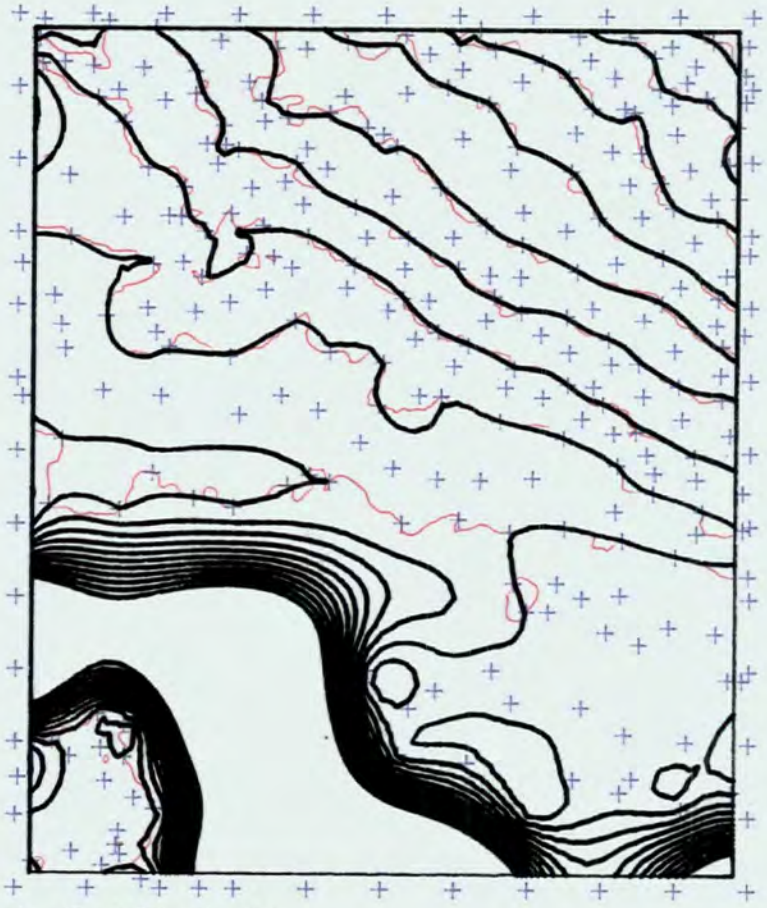
Thus a pointwise quadratic distance-weighted average of surrounding data points is used within MINCURV at this stage.

It should thus be impossible to derive extreme preliminary values, and accuracies obtained by similar methods (e.g. SURFACE II GRAPHICS) would be expected.

Although a minimal error may be generated in the first stage, it is from the second iterative stage that the major error sources derive. Considerable distortion may occur when using the difference equations associated with the corner and boundary zone (outer two rows and columns). In this area, the off-grid equations are not applied, but instead the data points are moved to the nearest grid points and these values fixed to control the surrounding grid values during iteration (Swain, 1978). An error, directly in relation to the node/data distance, is thus generated. Within the internal block and on the other non-defined external nodes, error is also closely linked with the node/data distance and any error that is already extant. Since the iterative interpolation procedure involves generating new values from old and the majority of old values within this evaluation are interpolated and thus in error, most error will be accumulated in proportion to the node/data point distance. Hence, it is obvious that error and its escalation may be diminished, possibly quite rapidly, either by restricting the height range (for example by using FORV data rather than INCH data) or by increasing the point density.

While nodes generated in the first stage of the interpolation are within tolerable limits, Figures 6.16 and 6.17 illustrate some of the problems inherent in the second stage. It is obvious from both figures that there is a large

Figure 6.16



L-C MINCURV

L-R MINCURV

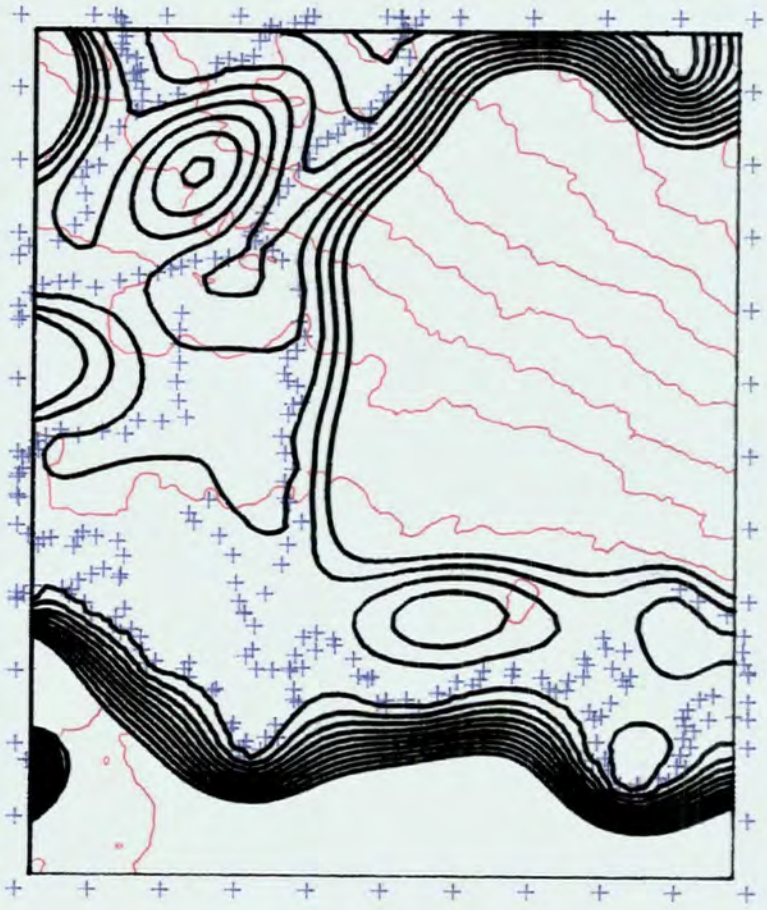
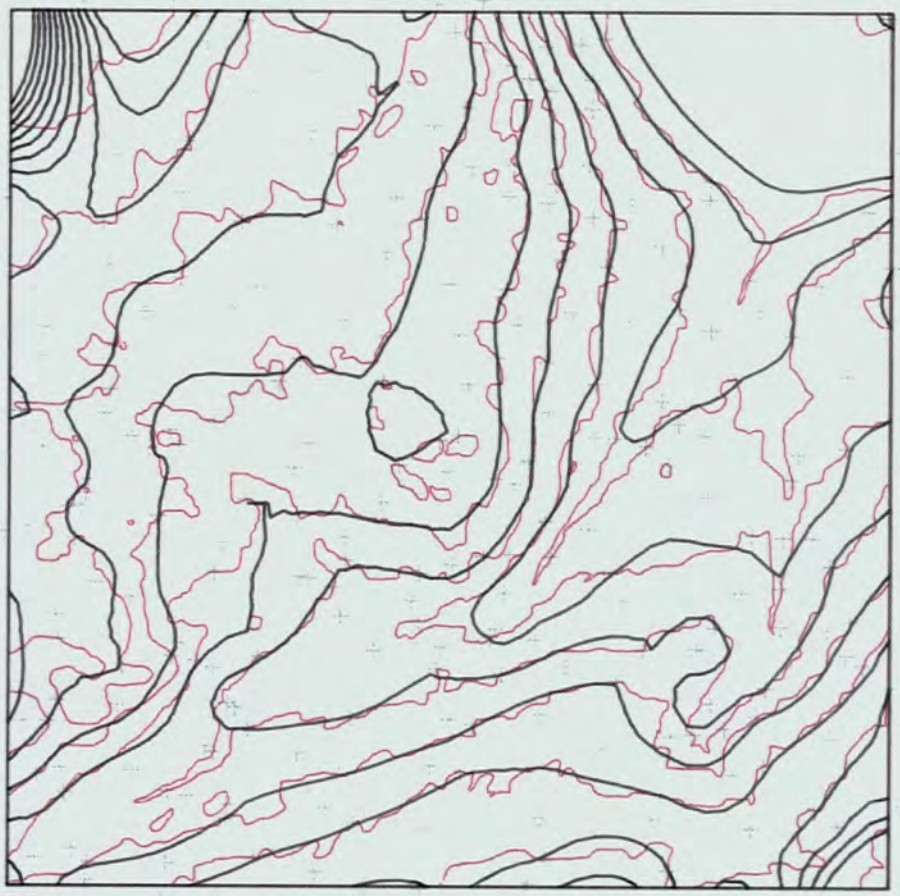
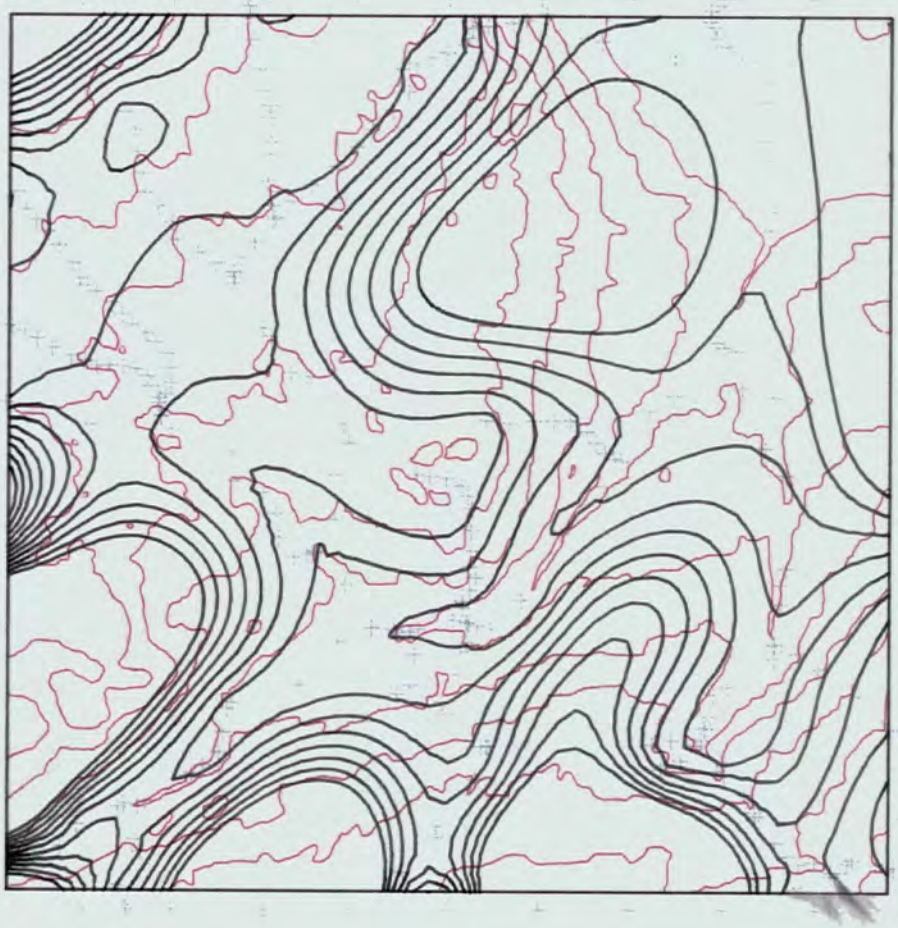


Figure 6.17



I-C MINCURV

I-R MINCURV



variance among different data sets within any model and thus error is not correlated with local surface morphology. It is apparent however that the data voids coincide exactly with the areas of isarithm voids (N.B. where the interpolated grid node lies outside the range of plotted isarithms). Therefore gross errors using this pseudo-splining technique are a direct result of data distribution. It has already been proposed that much of the error on the boundary zones has been introduced by the difference equations in conjunction with the node/data distances and, in relation to this, most major error sources occur on the boundaries. However, the boundary is not responsible for all error, as is illustrated by the plot of INCH-RIVER data. Finally, the results derived from INCH-CONTOUR data suggest that - where the data distribution allows - good results may be achieved for poorly distributed data, although this cannot be directly correlated to the NNS, as LIAN-CONTOUR data produce a higher value.

6.2.7 KRIGE

> "In theory, no other method of grid generation can produce more accurate estimates of the form of the mapped surface ... even with naive estimates of the parameters, kriging will do no worse than arbitrary estimating procedures such as those in 'GRID'" (Sampson, 1978, 111).

Examining all the SURFACE II GRAPHICS statistics (Table 6.3 to 6.7) and Table 6.13 (KRIGE), the interpolating ability of KRIGE is seen to be vastly superior to any of the other SURFACE II GRAPHICS methods in terms of accuracy. KRIGE, however, does not interpolate more accurate estimates than MULTI (Table 6.16) although it does produce marginally

DS	%BAD	ORIGINAL		TRIMMED		TRIMMED ERROR				TRIMMED ABSOLUTE ERROR			CORRELATIONS			
		MIN	MAX	MIN	MAX	MEAN	MEDIAN	SKEWNESS	KURTOSIS	MEAN	MEDIAN	STANDARD	SURFACE	NEAR	NONN	GRADIENT
IB	0.67	-16.94	20.35	-9.06	9.02	0.02	0.17	-0.13	2.81	2.85	2.31	3.59	0.997	0.251	-0.174	0.042
IC	1.25	-18.80	16.91	-7.89	7.89	-0.28	-0.35	0.09	2.77	2.51	2.06	3.14	0.997	0.135	-0.008	0.009
IG	1.04	-22.46	16.33	-7.84	7.81	-0.06	-0.22	0.12	3.59	1.89	1.39	2.60	0.998	0.523	-0.446	0.110
IR	8.57	-40.69	16.89	-22.61	16.50	-6.02	-4.55	-0.93	4.06	6.52	4.64	8.86	0.991	0.415	-0.267	0.270
IS	1.17	-18.33	27.03	-8.35	8.32	0.70	0.69	-0.07	2.93	2.54	2.06	3.22	0.997	0.155	0.046	0.160
LB	1.82	-6.89	6.64	-3.22	3.19	-0.69	-0.69	0.27	3.04	1.10	0.95	1.35	0.999	0.200	-0.183	-0.004
LC	1.86	-6.85	5.79	-2.67	2.79	-0.26	-0.25	-0.09	2.86	0.95	0.81	1.17	0.999	0.058	-0.096	0.137
LG	1.42	-7.46	5.30	-4.68	4.59	-0.09	-0.10	-0.72	5.96	0.70	0.49	0.94	0.999	0.438	-0.130	0.191
LR	7.30	-11.02	8.98	-5.59	5.59	-0.03	-0.26	0.44	3.34	1.52	1.17	2.00	0.996	0.409	-0.329	0.299
LS	2.35	-4.97	6.39	-2.57	2.56	-0.30	-0.35	0.33	2.89	0.85	0.74	1.06	0.999	0.112	-0.037	0.049
FB	0.00	-4.81	10.51	-4.81	4.81	0.33	0.48	0.88	4.38	0.98	0.66	1.39	0.967	0.269	-0.177	0.263
FC	0.02	-6.72	6.09	-2.09	2.10	0.14	0.13	0.05	2.88	0.61	0.51	0.77	0.990	0.066	0.140	0.113
FG	0.00	-10.44	4.50	-2.29	2.27	0.13	0.10	-0.01	4.67	0.47	0.32	0.66	0.992	0.309	-0.054	0.263
FS	0.02	-8.52	7.54	-3.51	3.56	0.17	0.06	0.59	4.30	0.74	0.52	1.02	0.981	0.149	0.007	0.193

Table 6.13a Primary results of KRIGE

DS	IN-NEAR			IN-NONN			IN-GRAD			OUT-NEAR			OUT-NONN			OUT-GRAD		
	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN
IB	0.95	174.83	50.63	0	14	3.79	0.12	83.55	19.24	9.01	172.25	68.13	0	10	2.67	1.42	77.96	21.83
IC	0.83	182.47	44.16	0	10	3.29	0.12	83.55	19.05	12.27	134.16	47.22	0	9	3.92	2.90	83.00	24.38
IG	0.00	42.42	32.11	4	6	4.21	0.17	83.00	19.11	30.00	42.42	38.02	4	5	4.02	0.12	83.55	24.19
IR	0.91	306.41	69.36	0	20	3.65	0.17	83.55	18.86	8.07	275.63	116.69	0	9	0.86	0.12	78.75	22.86
IS	0.31	150.02	46.25	0	12	3.32	0.12	83.55	19.00	6.85	102.74	54.97	0	8	3.15	2.67	77.96	25.14
LB	0.31	165.00	39.02	0	23	3.22	0.00	41.50	10.52	5.64	150.00	53.66	0	21	2.07	0.56	30.24	8.76
LC	0.00	133.99	28.46	0	11	3.26	0.00	41.50	10.42	4.23	30.00	17.47	2	4	3.29	2.30	6.85	4.29
LG	0.00	42.42	19.72	1	6	3.32	0.00	41.50	10.46	15.00	42.42	24.41	1	4	3.02	0.56	35.99	8.02
LR	0.31	226.99	53.50	0	19	3.56	0.00	41.50	10.19	4.38	226.99	88.67	0	12	0.98	0.00	39.77	11.88
LS	0.21	112.92	27.66	0	11	3.01	0.00	41.50	10.35	4.86	77.77	28.05	0	9	3.43	0.56	36.63	11.01
FB	0.21	128.80	27.47	0	18	4.79	0.00	84.47	8.52	12.85	129.47	64.77	0	12	1.69	0.13	61.88	10.27
FC	0.40	70.00	21.61	0	17	4.44	0.00	84.47	8.10	2.13	43.72	17.49	1	16	8.47	0.38	61.88	18.00
FG	0.00	41.22	16.22	2	7	4.77	0.00	84.47	7.97	10.00	40.00	19.97	2	7	4.43	1.07	74.47	20.58
FS	0.00	84.33	24.58	0	17	4.59	0.00	84.47	8.14	0.00	55.25	25.69	0	13	5.69	0.45	61.88	17.33

Table 6.13b Secondary results of KRIGE

superior results to the second most accurate form of interpolation studied - SACM (Table 6.14). However, KRIGE is inferior to many of these other methods as a result of its poor ability to interpolate in extensive data voids. The SURFACE II GRAPHICS manual (Sampson, 1978) suggests that this may be a factor of the run parameters selected, although many combinations were examined.

Kriging was accessed within SURFACE II GRAPHICS and thus the ORIGINAL MIN/MAX errors are artificially low due to certain nodes not being evaluated (see 6.2.1). This is fairly widespread (see bracketed values) and, if the magnitudes of the MIN/MAX values and the %BAD statistics are examined, it is apparent that only a few INCH-RIVER nodes were actually interpolated outside the arbitrary -40.0 to +40.0 'extreme error' range. The consistency of (on average) approximately 2% of the nodes not being evaluated, due to insufficient data points being found by the OCTANT search, is worrying. The conclusion must be drawn that the OCTANT search, as used in SURFACE II GRAPHICS for kriging, is too rigid. However, this cannot be directly associated to NNS, as the %BAD/NNS rank correlation is an insignificant 0.39, and no other significant characteristic associated correlations exist. This is associated with the high within-model consistency of %BAD, which suggests that, with the exclusion of exceptionally low NNS (say less than 1.0) data sets, each data set within any particular model will have a similar number of nodes which are not interpolated (INCH - 1.0, LIAN - 1.8 and FORV - 0.01).

Having removed many of the problem areas by not attempting

to interpolate the nodes, it is not surprising that the mean magnitude ORIGINAL MIN/MAX value is low. However a value of 12.29 which subsequently transfers into a TRIMMED MIN/MAX mean magnitude of 6.01 (GINOSURF - 7.93 and OCTANT 7.89) does show the power of this method. In addition to the exceptionally low mean values, there is (with the exclusion of RIVER data) a low within-model variance and the MIN and MAX values are well-balanced, suggesting an unbiassed interpolator. The other precision descriptor, KURTOSIS, displays a surprisingly high mean (3.61). While this is not as high as the more specialised pointwise methods (GPCP, GINOSURF and GHOT), it is greater than the basic weighted averaging methods. This is surprising since SKEWNESS magnitude and the TRIMMED error range are smaller than for the specialised methods, and should therefore decrease KURTOSIS. Therefore, errors are not only confined to a small range, but the distribution is noticeably peaked at the MEAN.

The high precision of the error distribution is supported by the similarity of the two measures of central tendency - MEAN and MEDIAN. With the exception of a few data sets (INCH-RIVER, INCH-SCATTER and LIAN-BREAKLINE), there is good within-model systematic error consistency. These measures (MEAN and MEDIAN) are also of low magnitude, underlining the low bias found in the ORIGINAL MIN/MAX statistics.

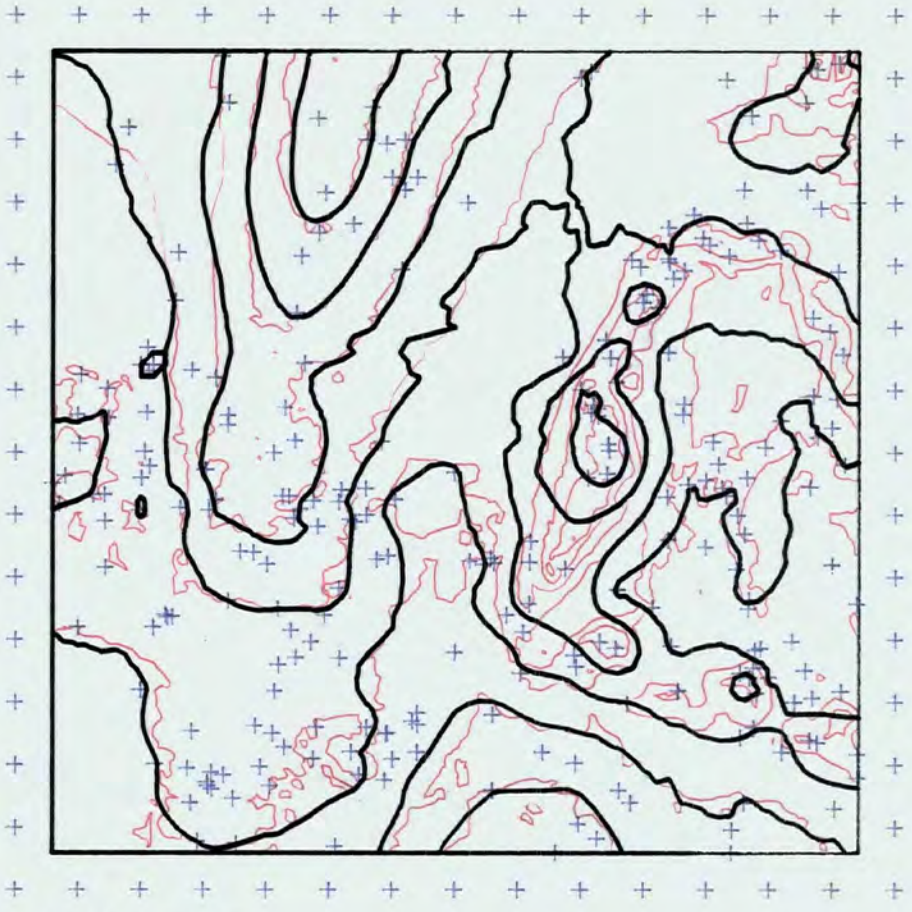
From the previous discussion, it is not surprising that the resultant within-model accuracies are remarkably consistent, and that they are superior to any of the previous methods (mean STANDARD ERROR - 2.27 as opposed to GINOSURF - 2.63).

While the accuracies of the data sets producing the poorest previous accuracies have obviously increased most rapidly (eg. INCH-BREAKLINE, LIAN-CONTOUR and FORV-SCATTER) the poor distribution of RIVER data still poses problems of accuracy. This stresses the fact that, even for KRIGE to be successful, data sets must have a NNS greater than about 1.0.

Importantly, there are no significant error-related correlations. Low mean values of 0.249 (ERROR/NEAR), -0.122 (ERROR/NONN) and 0.150 (ERROR/GRAD) underline the fact that this method of interpolation involves estimating local surface autocorrelation. The IN and OUT statistics confirm this, since most IN and OUT values are very similar. This suggests that the extreme errors and the nodes not interpolated may occur in a variety of situations which are not related to the nearness of data points to the nodes being interpolated, the number of data points close to each node being interpolated, or the gradient in the local vicinity of the grid nodes. The two data sets with NNS values of under 1.0 are exceptions and show a sharp contrast between the IN and OUT NONN-MEANs, suggesting that this is a factor crucial to poor interpolation for the low NNS data sets.

The benefits of KRIGE are illustrated in Figures 6.18 and 6.19. In general, excluding the %BAD - not interpolated areas, the surface morphology has been well-modelled from what are the most troublesome data sets. However, serious doubts must be raised about any procedure which fails to interpolate grid nodes so close to such dense clusters of data points (bottom-left of LIAN-RIVER and INCH-RIVER).

Figure 6.18



F-S KRIGE

L-R KRIGE

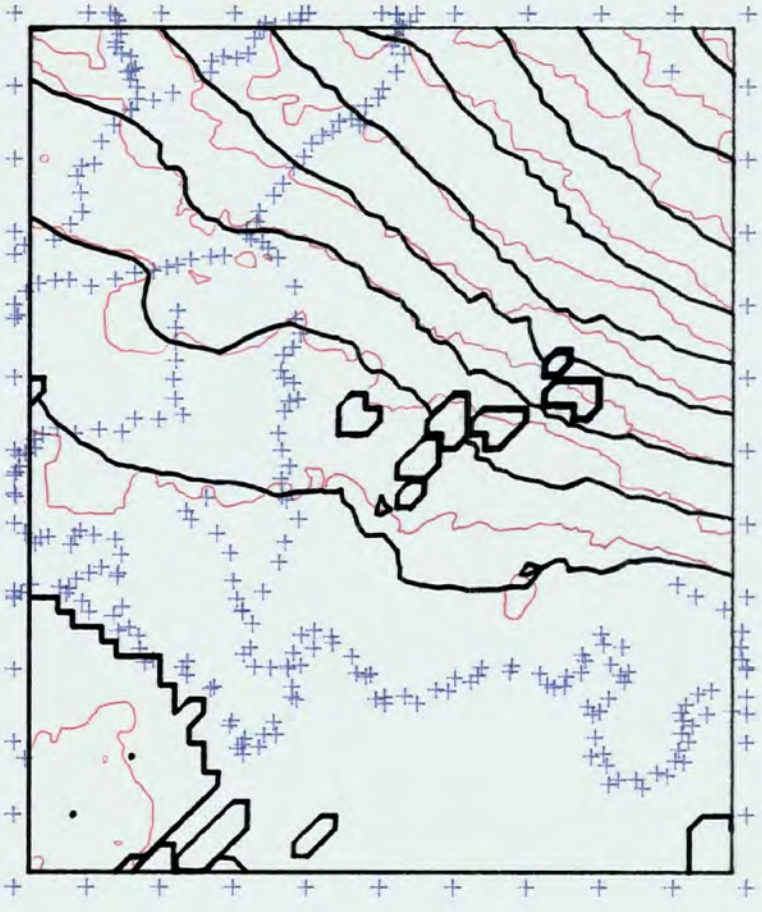
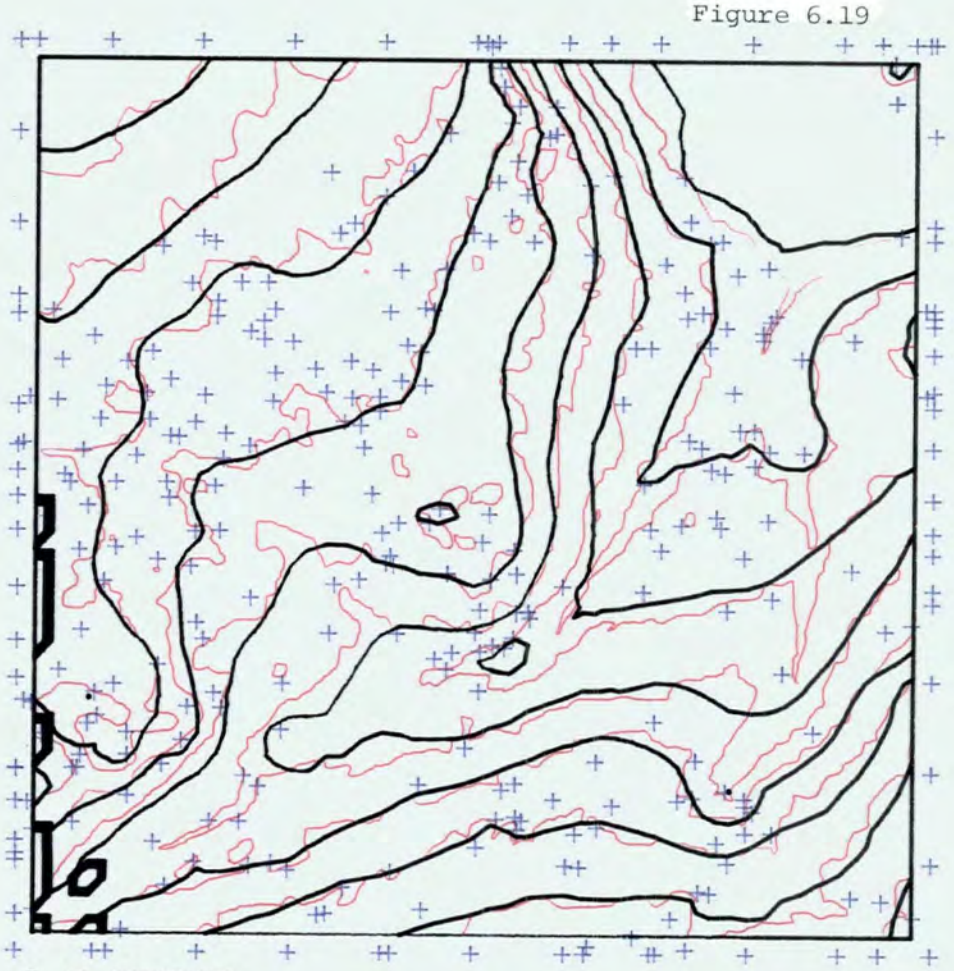
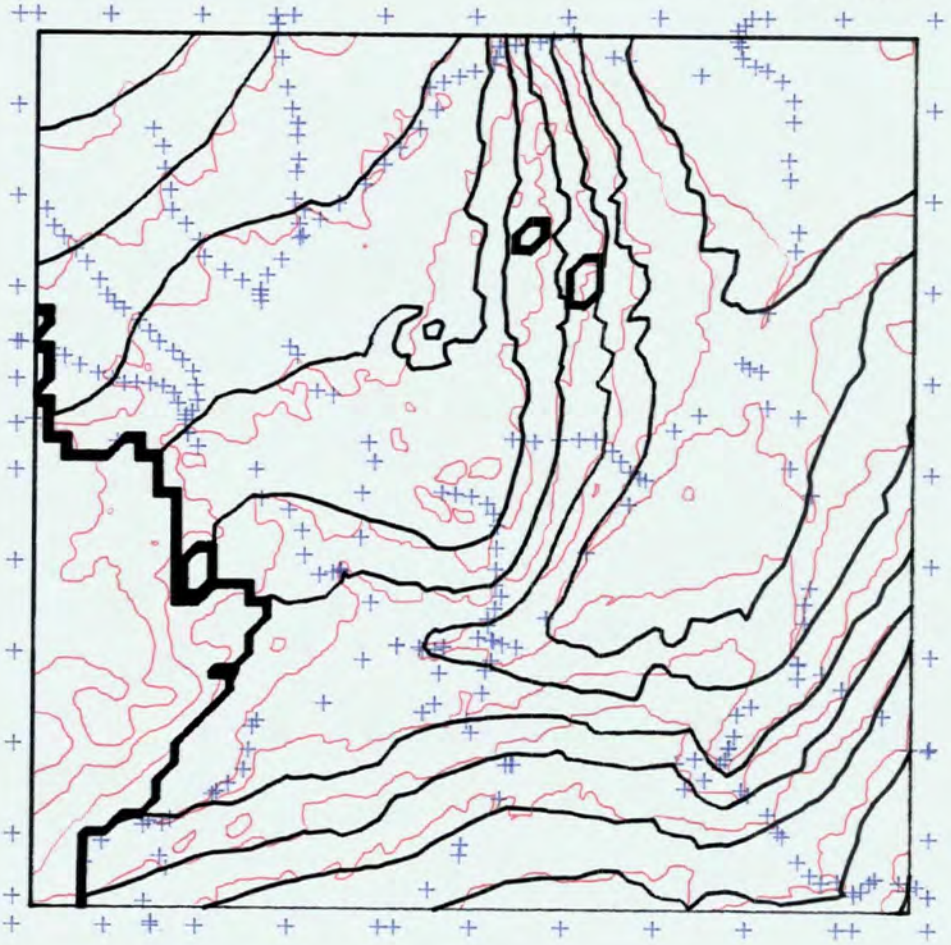


Figure 6.19



I-B KRIGE

I-R KRIGE



6.3 PATCHWISE METHODS

6.3.1 SACM

SACM derives its accuracy from its two-pass system of interpolation. First, grid cells containing data points are evaluated by a patchwise linear surface constrained to pass through the data points and fitted to 8 points found by a local octant search. This should eliminate interpolation extremes and relate the nodes of the cell to the whole surrounding local environment rather than a small sector of that environment (if a nearest 'n' point search was used). The second pass, by working outwards from previously derived grid nodes, ensures that derived points will almost certainly be reasonable and that a smooth transition will occur from one data point environment to another.

Examining the statistical results of interpolation, Table 6.14, the package is seen generally to be successful. The total %BAD statistic is the lowest encountered for any of the interpolation methods investigated suggesting a generally safe interpolator. Alternatively, the ORIGINAL MIN/MAX values are marginally and consistently greater than the comparable KRIGE values (mean magnitude 14.13 - SACM, 12.29 - KRIGE). This is to be expected, since there are many KRIGE nodes in data voids not interpolated. However, the MIN/MAX values are superior to all other methods, with the exclusion of MULTI, supporting the superior global accuracy and consistency of this method.

DS	%BAD	ORIGINAL		TRIMMED		TRIMMED ERROR				TRIMMED ABSOLUTE ERROR			CORRELATIONS			
		MIN	MAX	MIN	MAX	MEAN	MEDIAN	SKEWNESS	KURTOSIS	MEAN	MEDIAN	STANDARD	SURFACE	NEAR	NONN	GRADIENT
IB	0.00	-17.42	38.33	-9.85	9.97	0.20	0.14	-0.05	2.86	3.02	2.52	3.84	0.996	0.202	-0.109	0.045
IC	0.00	-20.02	18.27	-9.05	8.97	-0.19	-0.20	0.06	2.93	2.71	2.18	3.43	0.997	0.123	0.043	0.038
IG	0.00	-16.80	22.33	-12.89	14.27	0.06	-0.48	0.32	5.17	2.25	1.60	3.22	0.997	0.469	-0.407	0.165
IR	0.25	-45.28	16.66	-23.86	16.39	-6.07	-4.89	-0.69	3.39	6.56	4.94	8.60	0.991	0.394	-0.243	0.321
IS	0.00	-22.82	30.08	-9.98	10.06	0.88	0.71	0.01	2.89	3.01	2.39	3.82	0.996	0.155	0.057	0.123
LB	0.00	-6.81	9.39	-6.56	6.39	-0.61	-0.59	0.20	4.75	1.21	0.95	1.59	0.998	-0.040	0.072	0.130
LC	0.00	-6.35	10.11	-3.57	3.60	-0.03	-0.17	0.42	3.01	1.07	0.88	1.36	0.998	0.129	-0.105	-0.136
LG	0.00	-7.21	5.48	-4.97	4.80	-0.08	-0.13	-0.87	6.30	0.69	0.49	1.01	0.999	0.425	-0.121	0.188
LR	0.00	-13.05	7.17	-6.97	6.81	-0.13	-0.26	0.34	4.18	1.52	1.07	2.10	0.995	0.362	-0.241	0.292
LS	0.00	-5.39	8.93	-5.39	5.85	-0.22	-0.29	0.65	5.09	1.02	0.79	1.39	0.998	0.052	0.085	0.145
FB	0.00	-7.61	10.76	-5.05	5.25	0.56	0.24	0.40	5.14	1.05	0.71	1.52	0.960	0.284	-0.187	0.134
FC	0.00	-7.81	7.43	-2.84	2.89	0.21	0.15	0.14	3.39	0.74	0.56	0.97	0.984	0.107	0.099	0.102
FG	0.00	-10.85	7.56	-2.36	2.47	0.13	0.09	-0.05	4.93	0.49	0.33	0.69	0.991	0.338	-0.108	0.227
FS	0.00	-8.10	7.56	-4.31	4.32	0.20	0.09	0.36	4.71	0.88	0.59	1.25	0.971	0.132	0.039	0.212

Table 6.14a Primary results of SACM

DS	IN-NEAR			IN-NONN			IN-GRAD			OUT-NEAR			OUT-NONN			OUT-GRAD		
	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN
IB	0.95	174.83	50.77	0	14	3.77	0.12	83.55	19.05	9.01	172.25	67.41	0	10	2.88	1.00	83.00	25.61
IC	0.83	182.47	44.19	0	10	3.30	0.12	83.55	19.02	12.27	112.66	47.36	0	9	3.93	1.06	83.00	26.30
IG	0.00	42.42	32.41	4	6	4.19	0.12	83.55	19.35	30.00	42.42	38.27	4	4	4.00	8.68	60.17	27.39
IR	0.91	306.41	72.72	0	20	3.43	0.17	83.55	19.10	41.75	218.39	126.50	0	8	0.60	0.12	78.75	24.80
IS	0.31	150.02	46.35	0	12	3.31	0.12	83.55	18.92	17.05	100.97	55.05	0	9	3.43	2.90	65.16	28.08
LB	0.31	165.00	40.02	0	23	3.14	0.00	41.50	10.38	19.72	45.00	31.66	1	10	4.32	4.68	30.24	16.50
LC	0.00	133.99	28.02	0	11	3.31	0.00	41.50	10.56	7.11	109.00	36.05	0	10	2.39	0.00	40.00	7.38
LG	0.00	42.42	19.78	1	6	3.32	0.00	41.50	10.38	15.00	33.53	24.77	1	4	2.59	4.04	35.99	14.50
LR	0.31	226.99	57.80	0	19	3.27	0.00	41.50	10.29	16.30	142.88	54.30	0	12	1.86	4.14	39.77	16.13
LS	0.21	112.92	27.72	0	11	3.02	0.00	41.50	10.36	13.04	33.80	21.22	1	7	4.71	3.94	30.24	16.38
FB	0.21	129.47	29.33	0	18	4.64	0.00	84.47	8.60	31.61	36.05	33.83	3	3	3.00	3.09	3.88	3.47
FC	0.40	70.00	21.44	0	16	4.55	0.00	74.47	8.30	2.47	41.22	20.57	1	17	7.73	1.06	84.47	18.61
FG	0.00	41.22	16.21	2	7	4.76	0.00	84.47	7.93	10.00	40.00	20.22	2	7	4.47	1.07	74.47	21.50
FS	0.00	84.33	24.57	0	17	4.60	0.00	84.47	8.26	7.42	53.38	26.11	0	12	5.60	0.90	61.88	17.33

Table 6.14b Secondary results of SACM

Surprisingly, precision is not as high as might be expected (TRIMMED mean magnitude - 7.52), although this should be considered in the context of the consistently high KURTOSIS (mean 4.20). Systematic error and SKEWNESS are of generally low magnitude and therefore this emphasises the fact that the method constrains the majority of error within narrow limits. Where good data distribution occurs, the simpler first-stage SACM function, with localised searching, will produce many highly accurate results. Elsewhere, the surface averaging functions will produce safe acceptable values - it is noteworthy that the SURFACE II GRAPHICS interpolation methods produce TRIMMED MIN/MAX values most similar to SACM.

Accuracy is also marginally poorer than for KRIGE. There is low within-model variance for all descriptors with, once again, the exception of results based on INCH-RIVER data. Like all other methods, SACM cannot cope with the deep incision of the INCH valleys, which produces a strong negative bias to the interpolation results. However, the perfect (1.0) STANDARD ERROR and ABS MEAN rank correlations with NNS support the previous discussion concerning the association between the two stages of interpolation, their resultant accuracy and NNS.

All error-related correlation coefficients are low, underlining the problems of deriving these correlations where low error exists, but also suggesting that causation of error may be due to a multiplicity of factors. This could be a result of the two-stage interpolation process, both stages having differing causes of error. In addition, mean correlations of ERROR/NEAR - 0.224, ERROR/NONN - -0.080 and

ERROR/GRAD - 0.142 are similar to those of KRIGE, underlining the similarity between similar accuracy methods.

The statistics which are superior to all other methods excluding MULTI are emphasised by the superior isarithms interpolated (Figures 6.20 and 6.21). The isarithms for the FORV data replicate the reference isarithms well, especially FORV-CONTOUR (for comparison with GINOSURF see Figure 6.13), although those from FORV-SCATTER also maintain the general shape of the surface. More significantly, INCH-CONTOUR and -RIVER data interpolate a grid which subsequently produces arguably the best isarithms for their respective data sets.

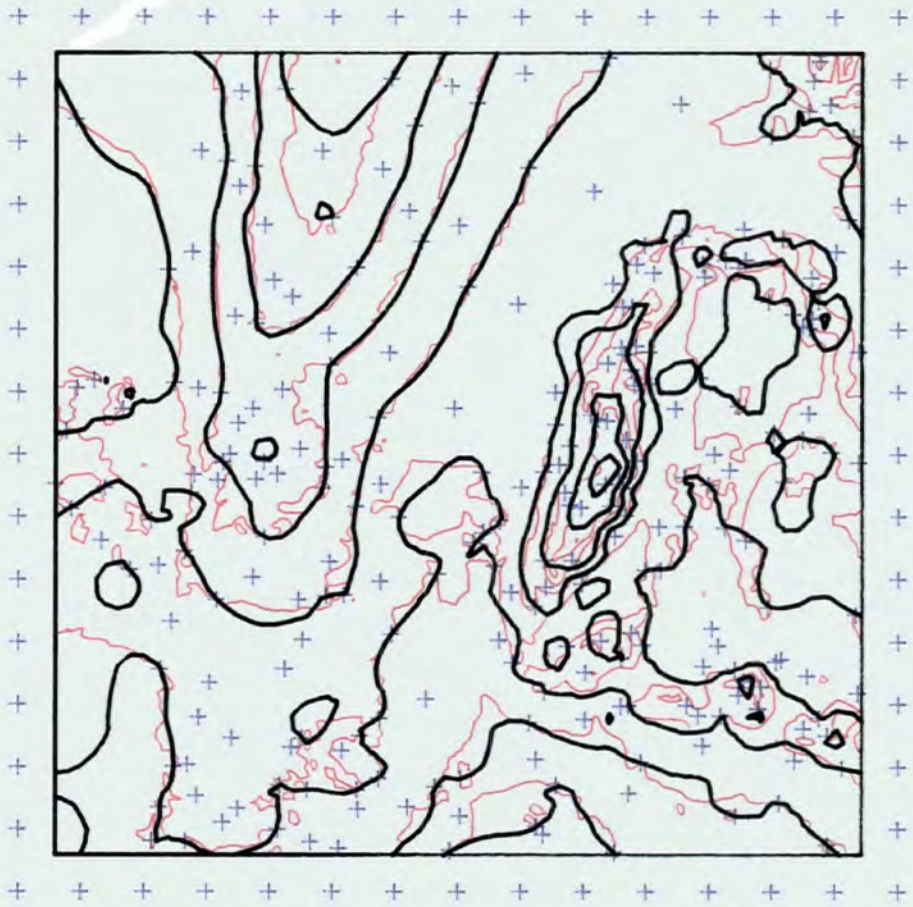
The success of this method must therefore not be measured solely in terms of specific accuracy, but also its consistency within individual map areas, data sets and models. Significantly, only INCH-RIVER produces a non-zero %BAD statistic. The robustness of the first-pass linear patchwise function, and the constraining of the second pass to build from the resultant interpolated nodes, ensure that no major error may occur.

6.3.2 BRAILE

When running BRAILE, four major parameters had to be supplied in addition to those defining grid and data sample size and co-ordinates. The size of the patches were defined:

a. in terms of the size of the patch of grid nodes being interpolated - set to 4x4;

Figure 6.20



F-C SACM

F-S SACM

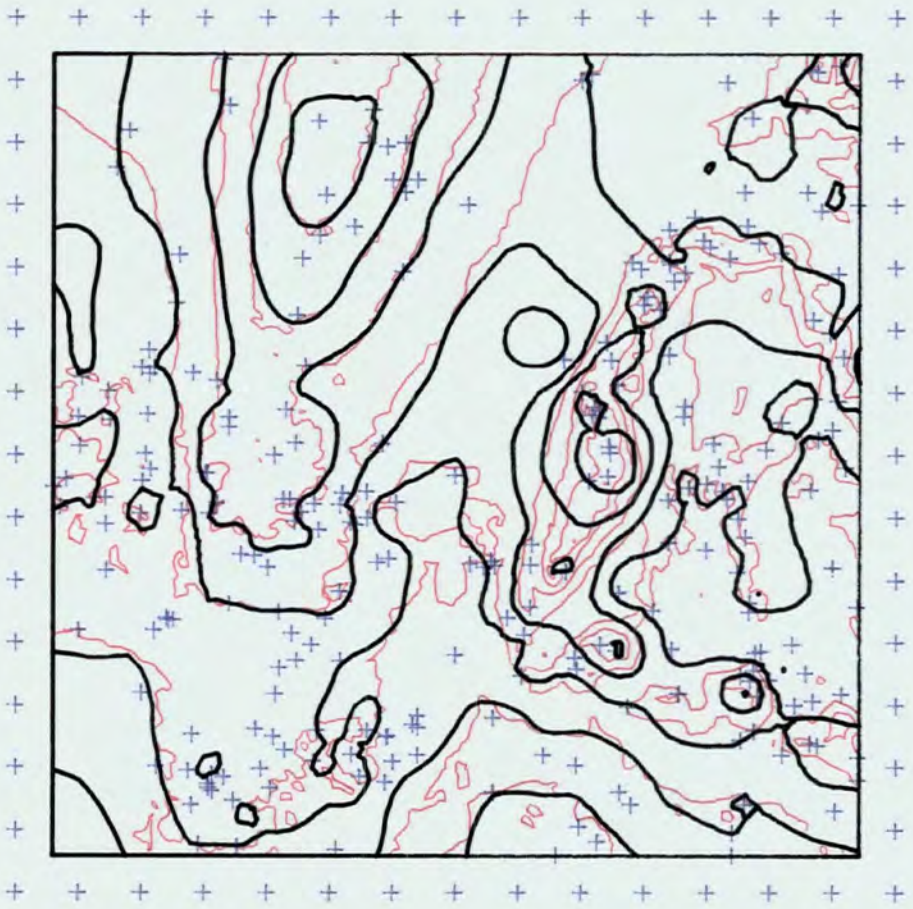
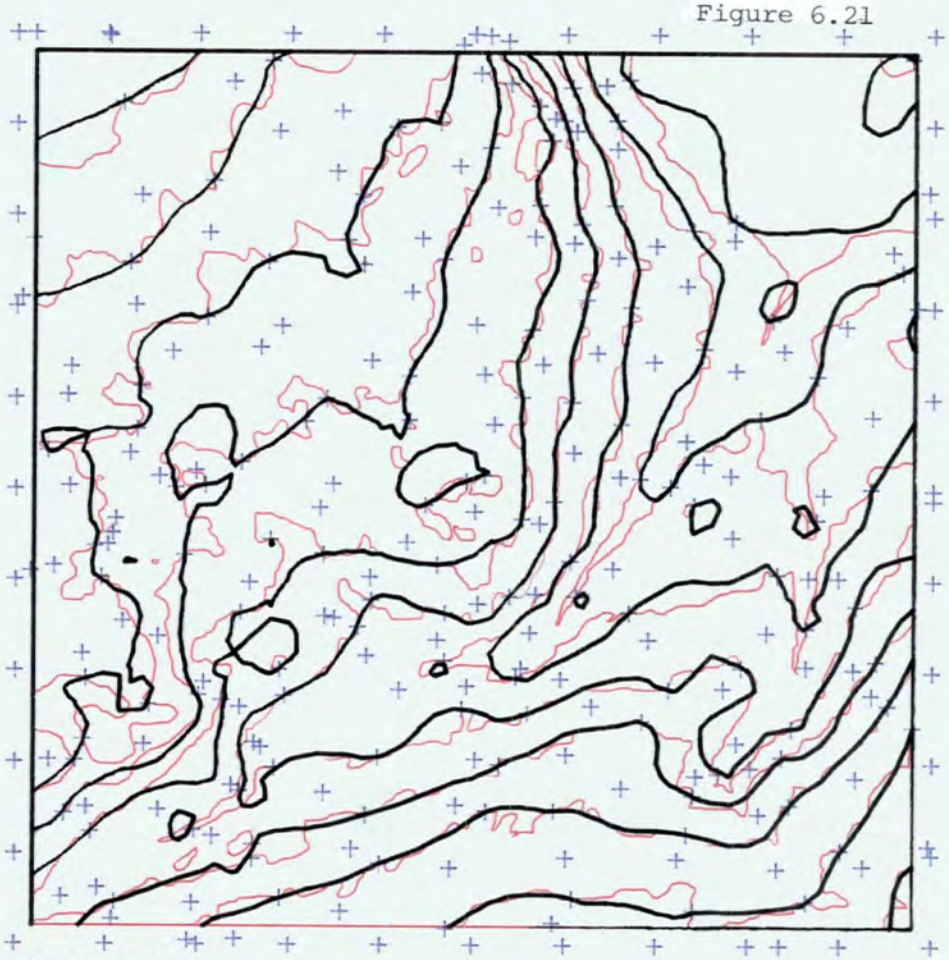
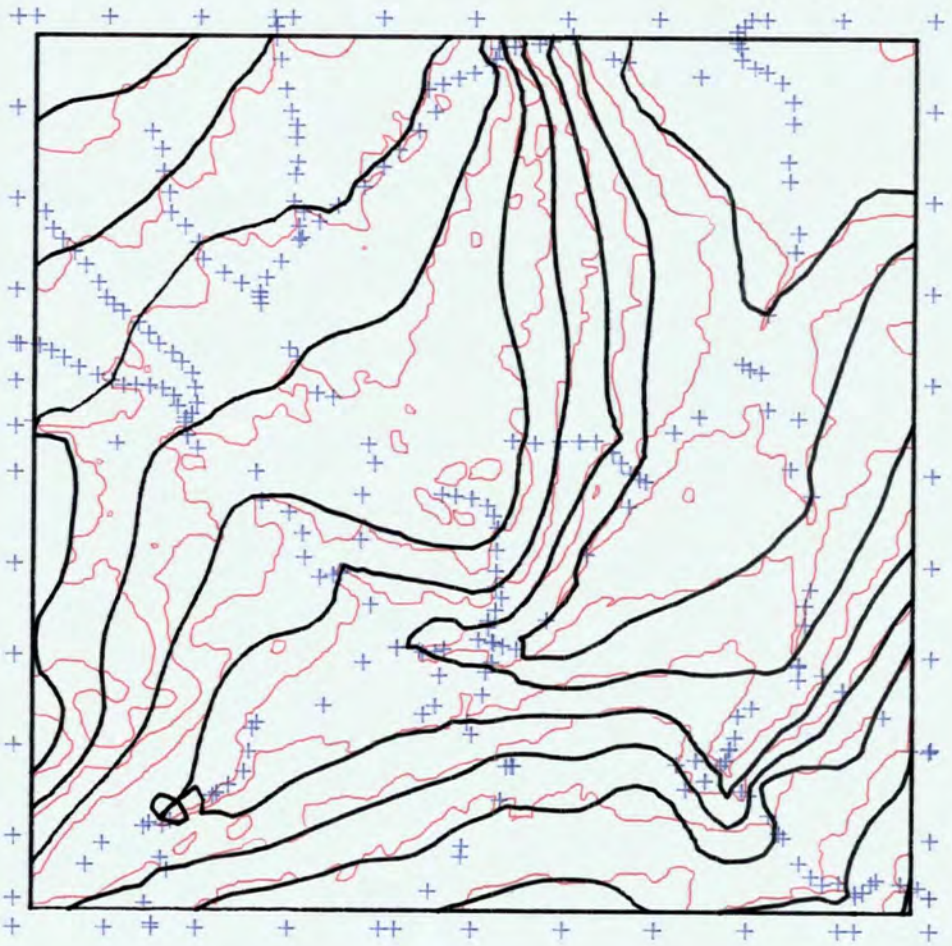


Figure 6.21



I-C SACM

I-R SACM



b. in terms of the search area of interest - set to 500x500m (INCH), 300x300m (LIAN) and 250x250m (FORV).

The maximum order of polynomial fit within each patch was set to third-order throughout and the minimum number of points required within each patch would equal the number of coefficients. Thus, if insufficient points were found for a third-order polynomial, the order would diminish until the number of coefficients was equivalent to the number of points - rather than in some previously discussed methods where the search radius was extended.

Overall, very few interpolated points were outside the required tolerance. The extreme errors from the BREAKLINE data are a direct result of poor interpolation at the edges - all the errors occur on the edge grid nodes and are therefore partly excusable. The gross errors from the RIVER data are outside the grid edge area and must be considered a direct result of the data. In general, the magnitude of the ORIGINAL ERROR MIN/MAX is greater than for most methods (see Table 6.15). The mean magnitude value of 23.80 (excluding the INCH-BREAKLINE MAX) is only surpassed by MINCURV, GPCP-TREND and the pointwise quadratic methods GINOSURF and GHOT. More importantly, with the exclusion of the very low NNS data sets, no significant correlation emerges between the ORIGINAL ERROR magnitude and any of the surface or data characteristics. This is an inherent aspect of the method, since a variety of functions and data redundancy are involved for solving each of the patches (and therefore each of the nodes) within each of the data sets. Each situation has its own associated causation

DS	%BAD	ORIGINAL		TRIMMED		TRIMMED ERROR				TRIMMED ABSOLUTE ERROR			CORRELATIONS			
		MIN	MAX	MIN	MAX	MEAN	MEDIAN	SKEWNESS	KURTOSIS	MEAN	MEDIAN	STANDARD	SURFACE	NEAR	NONN	GRADIENT
IB	0.37	-27.97	5355.70	-17.17	17.08	0.21	0.01	0.10	3.04	4.93	3.92	6.27	0.990	0.031	-0.051	0.175
IC	0.00	-34.91	30.19	-15.30	15.35	-0.59	-0.60	0.15	2.97	4.64	3.72	5.89	0.991	-0.046	0.088	0.104
IG	0.00	-34.11	33.38	-15.94	16.08	-0.16	-0.21	0.17	3.17	4.64	3.71	5.90	0.991	0.056	-0.069	0.180
IR	3.04	-160.36	39.25	-35.97	31.08	-7.61	-6.19	-0.56	3.89	9.26	6.90	12.34	0.977	0.362	-0.300	0.276
IS	0.00	-28.52	39.00	-16.08	16.08	0.16	0.02	0.02	2.89	4.98	4.05	6.39	0.989	-0.051	0.074	0.155
LB	0.00	-15.61	6.19	-4.35	4.35	0.31	0.34	-0.27	2.64	1.50	1.30	1.84	0.997	-0.032	0.060	0.109
LC	0.00	-7.93	7.50	-5.50	5.43	0.28	0.21	0.03	2.85	1.54	1.27	1.93	0.996	0.035	-0.015	-0.041
LG	0.00	-8.32	4.81	-3.00	3.00	0.05	0.03	-0.08	2.89	0.92	0.74	1.17	0.999	0.060	-0.044	0.259
LR	0.17	-51.00	14.07	-6.72	6.67	-0.60	-0.54	-0.13	3.32	1.81	1.32	2.39	0.994	0.312	-0.250	0.445
LS	0.00	-7.55	5.25	-3.32	3.32	0.03	-0.09	0.16	2.43	1.17	1.02	1.43	0.998	-0.028	0.046	0.068
FB	0.00	-10.13	26.13	-5.88	5.89	0.41	0.29	0.20	3.64	1.44	1.10	1.92	0.933	0.170	-0.132	0.137
FC	0.00	-8.63	7.10	-4.22	4.30	0.56	0.50	-0.05	2.94	1.26	1.00	1.61	0.957	-0.071	0.209	0.105
FG	0.00	-12.13	4.69	-3.39	3.39	0.27	0.28	-0.19	3.02	0.99	0.80	1.25	0.971	0.027	0.006	0.154
FS	0.00	-7.32	10.52	-4.75	4.75	0.57	0.42	0.26	3.31	1.22	0.91	1.60	0.959	-0.017	0.066	0.150

Table 6.15a Primary results of BRAILE

OS	IN-NEAR			IN-NONN			IN-GRAD			OUT-NEAR			OUT-NONN			OUT-GRAD		
	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN
IB	0.95	174.83	52.08	0	14	3.69	0.13	83.55	19.11	5.89	119.11	42.38	0	11	4.14	1.86	79.92	24.30
IC	0.83	182.47	44.88	0	10	3.27	0.12	83.55	19.00	1.92	99.33	34.24	0	8	4.47	2.11	79.92	26.64
IG	0.00	42.42	32.39	4	6	4.21	0.12	83.55	18.99	0.00	42.42	33.14	4	5	4.13	1.86	65.32	26.85
IR	0.91	306.41	71.24	0	20	3.47	0.17	83.55	19.05	60.00	291.54	154.94	0	2	0.12	0.12	68.89	25.66
IS	0.31	150.02	47.08	0	12	3.27	0.12	83.55	18.91	1.88	108.16	40.97	0	11	4.21	1.86	65.32	28.30
LB	0.31	165.00	40.47	0	23	3.05	0.00	41.50	10.38	1.50	105.28	30.97	0	23	4.89	0.50	36.63	10.88
LC	0.00	133.99	28.52	0	11	3.25	0.00	41.50	10.39	3.35	103.80	18.47	0	8	4.77	0.79	27.72	11.68
LG	0.00	42.42	19.72	1	6	3.34	0.00	41.50	10.14	0.00	42.42	21.63	1	5	3.02	2.68	39.77	15.30
LR	0.31	226.99	56.00	0	19	3.36	0.00	41.50	10.06	13.94	217.47	90.96	0	9	0.94	2.68	36.63	16.88
LS	0.21	112.92	27.99	0	11	2.97	0.00	41.50	10.31	2.47	85.99	22.02	0	9	4.22	0.74	36.63	12.17
FB	0.21	121.91	27.41	0	18	4.81	0.00	84.47	8.63	1.31	129.47	65.82	0	12	1.29	0.13	56.39	8.13
FC	0.40	70.00	21.57	0	17	4.48	0.00	84.47	8.43	2.47	43.78	18.47	1	15	7.65	0.56	56.39	11.98
FG	0.00	41.22	16.38	2	7	4.75	0.00	84.47	7.97	0.00	41.22	16.99	2	7	4.76	0.38	74.47	20.50
FS	0.00	84.33	24.41	0	17	4.67	0.00	84.47	8.48	0.00	62.22	28.89	0	12	4.19	0.13	56.39	10.89

Table 6.15 Secondary results of BRAILE

of error.

The precision descriptors show much similarity with those of SYMAP. The TRIMMED MIN/MAX values (mean magnitude value - 9.95) are only superior to MINCURV, GPCP-TREND, GHOT and SYMAP, although this method has the least within-model variance of the individual values. Greater consistency is also very evident within KURTOSIS. The mean KURTOSIS of 3.08 is one of the lowest means achieved and has low variance across the full range of data sets. This illustrates the near-normal distribution of error generated by BRAILE, especially if it is remembered that the TRIMMED limits of the error distribution are far apart, which would tend to marginally increased KURTOSIS. As with SYMAP, this emphasises that the error generation process of BRAILE must be considered a random process.

It is therefore not surprising that the error-related correlations are so low. The mean values for ERROR/NEAR (0.058), ERROR/NONN (0.013) and ERROR/GRAD (0.162) are some of the lowest encountered. In addition, there is considerable variance within the NEAR and NONN correlations. This is emphasised by the IN and OUT statistics. With the exclusion of a few very low NNS data sets, the IN and OUT versions of all NEAR and NONN statistics are comparable. However, the IN-GRAD MEAN and OUT-GRAD MEAN values show a weak yet consistent increase from IN to OUT. All the functions used within BRAILE define local surfaces. Where the local gradient is higher, the local surface functions are liable to be in error, with error generally related to the magnitude of the gradient. This is

supported by the low consistent ERROR/GRAD correlations.

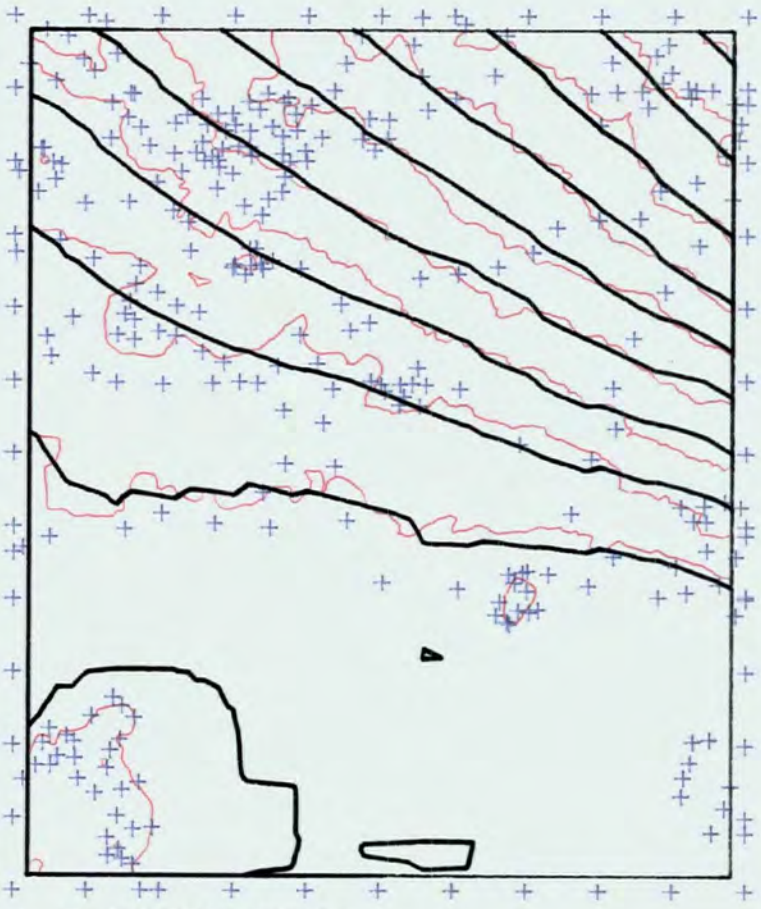
The other statistics support the previous theory. Excluding INCH-RIVER data, all MEAN and MEDIAN values have a magnitude less than 0.6 and generate mean values of 0.12 (MEAN) and 0.05 (MEDIAN). SKEWNESS values are also of low magnitude. These would suggest that the normally distributed random error is near centrally distributed about 0.0. The accuracy statistics reflect the characteristics of the precision and systematic error descriptors. In general, with the exclusion of RIVER data, there is low within-model variance, although the mean values are greater than for most other methods (ABS MEAN - 2.88, STANDARD ERROR - 3.71). This reflects the random error component associated with this method and is closely linked with the selection of such a large search area for each patch, which tends to smooth the final product.

The only accuracy/data characteristic correlation that seems important is that between accuracy and NNS (0.77). Clearly, where many points are found within a searched patch, the resultant interpolated points are constrained to lie within a reasonable tolerance of the true value. However, where the number of coefficients equals the number of points found, especially if the points lie in a small corner of the patch, the resultant fit in the area of interest surrounding the nodes will be poor. It is surprising that this potential correlation should not be reflected in ERROR/NEAR or ERROR/NONN correlations. This may be a result of the interpolation being patchwise and thus any correlation should be related to the whole patch, rather than to individual

points in the patch. Higher correlations might therefore be obtained if the patch size was reduced and the number of nodes evaluated in each patch was reduced. Similarly, further gross sources of error could be minimised by proportionally increasing the number of points required for a particular order of polynomial, since a greater averaging effect would be obtained.

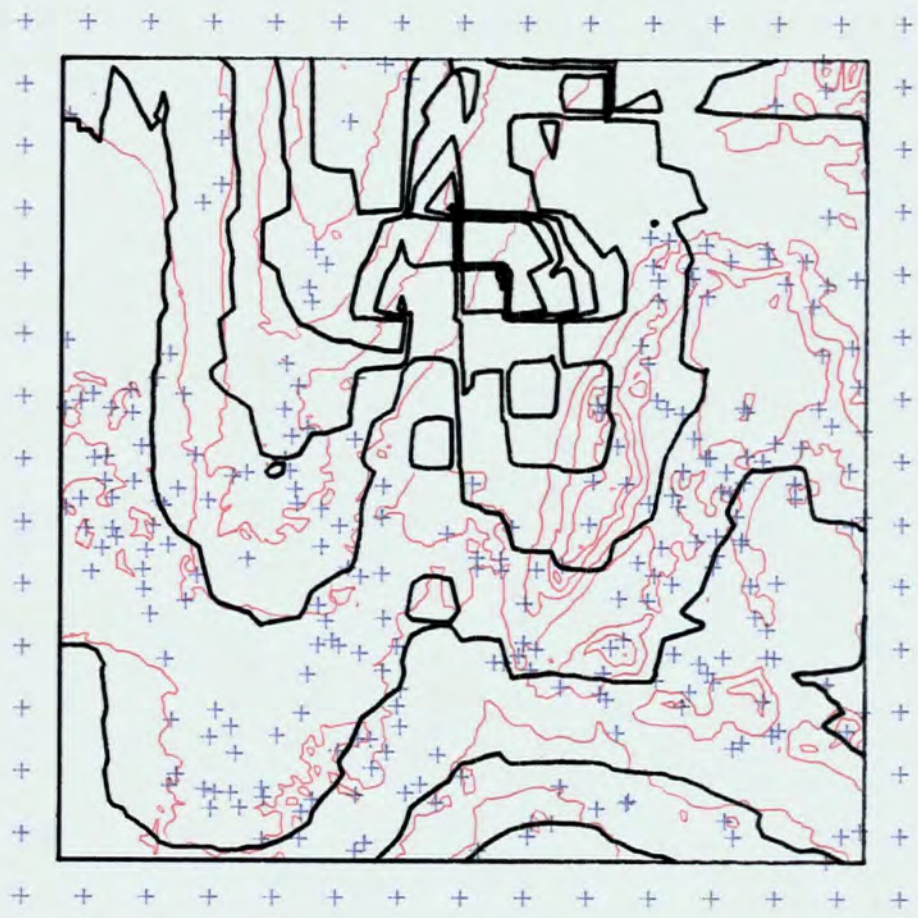
A major disadvantage of patchwise polynomials, especially without some form of splining at the edges of the patches (to ensure continuity between patches), is that - while their accuracy statistics may look reasonable - the morphological trueness of the final isarithms may be poor and distinctly segmented (see Figure 6.22). As with the similar effect produced by some pointwise methods (e.g. GPCP), this is caused by contiguous patches obtaining data points from different morphological features or parts of a feature; it occurs mainly in areas of data voids. For example (see Figure 6.23), if one patch includes points only at the top of a slope, a horizontal linear surface may result. If a contiguous patch includes points only at the bottom of a slope, a parallel but lower linear surface will result, with a cliff in between. This 'PATCHWISE EFFECT' varies with roughness and also the ratio of the wavelength of that roughness to the size of the search patch. Thus LIAN-BREAKLINE, with a lower NNS but smoother surface, produces a grid of good morphological trueness. FORV-BREAKLINE, with a higher NNS yet rougher surface (see vector dispersion analysis - Table 4.8), produces poor morphological trueness. The maximum possible patchwise effect is thus determined by the form of the data surface; however

Figure 6.22



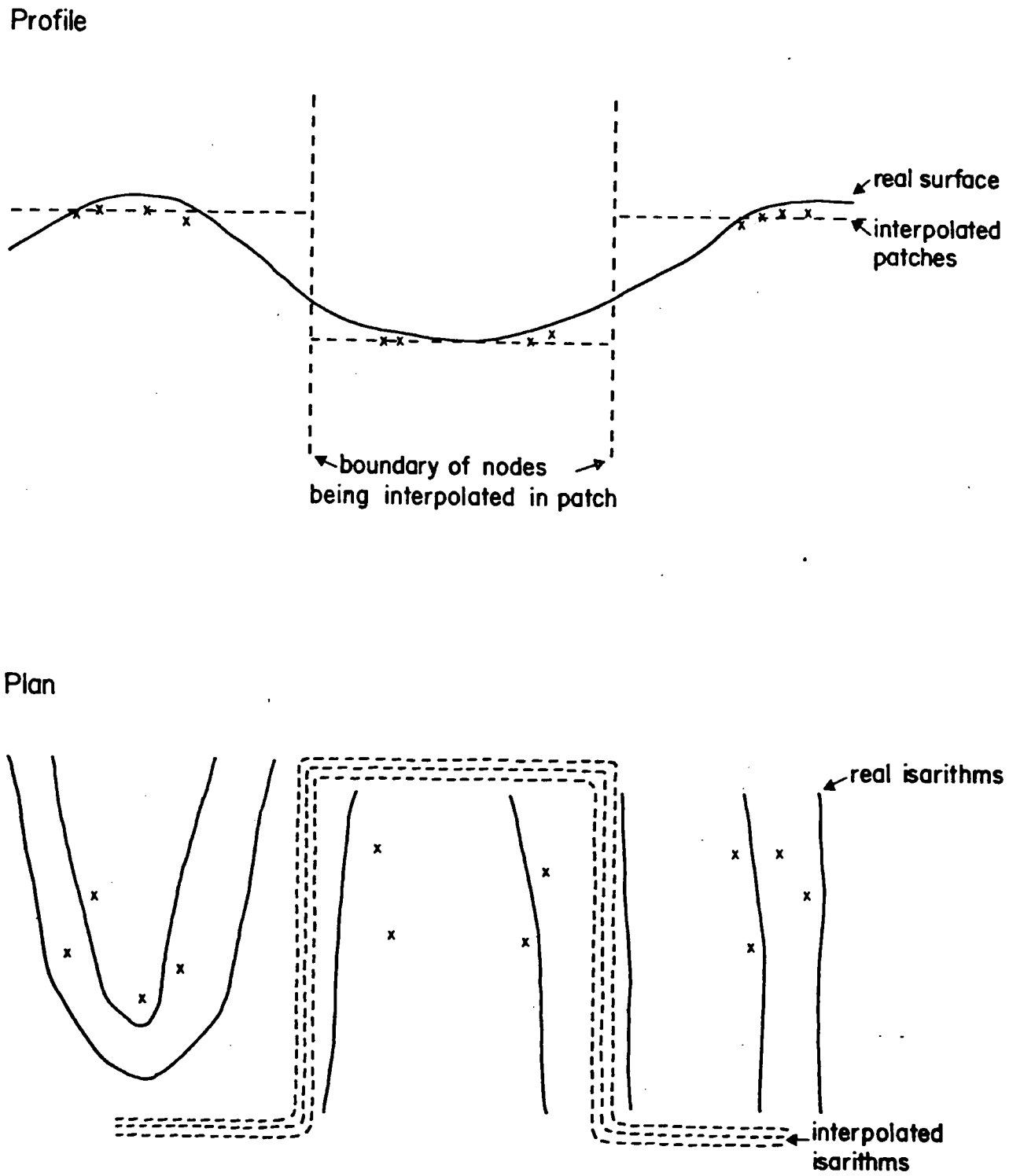
L-B PATCH.POLY

F-B PATCH.POLY



THE PATCHWISE EFFECT

Figure 6.23



its effect may be controlled by increasing the search patch, eventually generating a global fit.

6.4 HYBRID METHODS

6.4.1 MULTI

There is no doubt that the most statistically accurate results were produced by MULTI, the program which performs multiquadric analysis (see Table 6.16). While other methods perhaps generate occasionally superior results, MULTI consistently interpolates high accuracy results with few poorly estimated points (which only occurred with the INCH-RIVER data set).

The basic success of this method, then, lies in its consistency. In addition to a very low mean magnitude ORIGINAL MIN/MAX value (11.43), the individual values for many data sets are superior to the equivalent TRIMMED values for other methods. Excluding RIVER data, there is a very low within-model variance.

The TRIMMED MIN/MAX values behave similarly. The mean magnitude (5.37) is the lowest achieved for any method and, again with the exclusion of RIVER data, the individual values show a remarkably low within-model variance. Even with such a low distribution range, the KURTOSIS statistics (mean - 3.37) suggest a degree of peakedness. Thus, not only is error confined to a very narrow band but it has a strong central

DS	%BAD	ORIGINAL		TRIMMED		TRIMMED ERROR				TRIMMED ABSOLUTE ERROR			CORRELATIONS			
		MIN	MAX	MIN	MAX	MEAN	MEDIAN	SKEWNESS	KURTOSIS	MEAN	MEDIAN	STANDARD	SURFACE	NEAR	NONN	GRADIENT
IB	0.00	-13.22	20.30	-8.57	8.63	0.14	0.21	-0.05	2.71	2.73	2.26	3.40	0.997	0.249	-0.168	0.043
IC	0.00	-18.33	14.23	-7.68	7.63	-0.36	-0.46	0.10	2.71	2.52	2.10	3.17	0.997	0.157	-0.034	-0.004
IG	0.00	-15.29	16.57	-7.10	7.06	-0.06	-0.07	0.11	3.43	1.85	1.36	2.52	0.998	0.535	-0.453	0.111
IR	0.75	-50.36	9.13	-23.92	8.13	-7.00	-5.50	-0.99	3.52	7.09	5.55	9.50	0.990	0.510	-0.348	0.284
IS	0.00	-16.42	28.97	-8.14	8.15	0.63	0.64	-0.10	2.85	2.50	2.09	3.14	0.998	0.162	0.046	0.136
LB	0.00	-6.40	5.52	-2.79	2.75	-0.76	-0.79	0.44	3.22	1.02	0.94	1.25	0.999	0.268	-0.201	-0.044
LC	0.00	-6.13	5.56	-2.88	2.86	-0.11	-0.22	0.28	2.68	0.94	0.82	1.14	0.999	0.041	-0.020	-0.090
LG	0.00	-7.27	4.60	-2.18	2.15	-0.02	-0.11	0.00	3.32	0.57	0.45	0.76	0.999	0.477	-0.089	0.137
LR	0.00	-10.96	5.29	-4.39	4.29	-0.63	-0.60	0.10	3.21	1.29	1.05	1.64	0.998	0.282	-0.232	0.247
LS	0.00	-4.93	4.85	-2.46	2.44	-0.28	-0.32	0.28	2.82	0.82	0.70	1.01	0.999	0.097	-0.028	0.054
FB	0.00	-4.82	10.22	-4.82	5.14	0.26	0.18	0.94	4.47	1.02	0.62	1.48	0.964	0.320	-0.209	0.278
FC	0.00	-6.63	6.09	-2.14	2.14	0.17	0.15	0.03	3.14	0.59	0.46	0.75	0.991	-0.019	0.238	0.176
FG	0.00	-7.46	4.46	-2.35	2.29	0.13	0.10	0.01	4.67	0.47	0.31	0.66	0.992	0.305	-0.046	0.260
FS	0.00	-8.54	7.47	-3.52	3.73	0.28	0.13	0.60	4.35	0.76	0.51	1.06	0.980	0.145	0.007	0.212

Table 6.16a Primary results of MULTI

DS	IN-NEAR			IN-NONN			IN-GRAD			OUT-NEAR			OUT-NONN			OUT-GRAD		
	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN
IB	0.95	178.83	50.96	0	14	3.79	0.12	83.55	19.13	8.05	156.77	63.80	0	9	2.57	1.77	83.00	24.13
IC	0.83	182.47	44.16	0	10	3.29	0.12	83.55	19.05	12.27	99.33	47.85	0	9	4.05	2.90	83.00	25.50
IG	0.00	42.42	32.16	4	6	4.21	0.17	83.00	19.07	30.00	42.42	37.46	4	5	4.02	0.12	83.55	25.33
IR	0.91	306.41	72.00	0	20	3.44	0.17	83.55	19.08	51.11	291.54	140.50	0	4	0.33	0.12	78.75	25.10
IS	0.31	150.02	46.30	0	12	3.32	0.12	83.55	18.94	17.05	102.74	56.16	0	8	3.19	2.90	77.96	27.72
LB	0.31	164.44	39.39	0	23	3.15	0.00	41.50	10.27	5.64	165.00	51.36	0	23	2.82	1.23	32.25	12.76
LC	0.00	133.99	28.67	0	11	3.27	0.00	41.50	10.44	3.75	104.30	23.66	0	10	3.22	0.00	40.00	9.52
LG	0.00	42.42	19.52	1	6	3.35	0.00	40.00	10.09	15.00	42.42	25.49	1	4	2.89	2.23	41.50	16.39
LR	0.31	226.99	56.61	0	19	3.31	0.00	41.50	10.09	4.38	221.52	79.27	0	14	1.84	2.69	39.77	16.32
LS	0.21	112.92	27.61	0	11	3.00	0.00	41.50	10.19	4.86	72.99	29.07	0	9	3.82	0.80	36.63	14.27
FB	0.21	129.05	27.33	0	18	4.81	0.00	84.47	8.60	15.10	129.47	67.39	0	11	1.40	0.13	61.88	8.79
FC	0.40	70.00	21.60	0	17	4.44	0.00	84.47	8.13	2.13	41.97	17.85	1	16	8.39	0.38	61.88	17.69
FG	0.00	41.22	16.22	2	7	4.77	0.00	84.47	7.97	10.00	40.00	20.03	2	7	4.43	1.07	74.47	20.63
FS	0.00	84.33	24.49	0	17	4.60	0.00	84.47	8.19	6.54	58.88	27.30	0	13	5.43	0.45	55.25	16.35

Table 6.16b Secondary results of MULTI

tendency around a MEAN which, like the MEDIAN, has a generally low magnitude. The low error results have little correlation with any of the specified features (mean ERROR/NONN - -0.110, mean ERROR/GRAD - 0.129), although relatively strong ERROR/NEAR correlations for all the GRID and RIVER data sets result in a weak, yet noticeable mean correlation (0.252).

The accuracy descriptors show consistency with the previous error statistics. The mean ABS MEAN (1.73) and mean STANDARD ERROR (2.25) are not as distinct as the precision statistics, and have a similar magnitude to those of KRIGE (1.73 and 2.27) and to a lesser extent SACM (1.88 and 2.49). However, they are distinctly superior to those from any other method and MULTI is more reliable (see ORIGINAL MIN/MAX). With the exclusion of RIVER data, there is a low within-model variance.

Irrespective of data distribution, MULTI produces good results, which is a direct result of each data point being honoured. Each data point has a quadric centred on itself and fitted through all other data points. Thus, as has been shown with GINOSURF, good estimation occurs when a quadric is fitted over a broad area of interest, including a wide dispersion of points.

Yet, while the error statistics may be superior, the resultant isarithms generated from MULTI interpolated grids do not model the surface as well as those from SACM-interpolated grids. This may be considered to be a direct result of using a more global method of interpolation, which must result in greater smoothing, since every data point affects the interpolation of every grid node. This is illustrated in

Figure 6.24. Both FORV-CONTOUR (see Figure 6.20) and INCH-RIVER (Figure 6.21) produce more rounded isarithms than their SACM equivalents, in addition to minor features being poorly defined - for example the bottom-right river valley modelled by INCH-RIVER data and the sand dune summits in the centre- and lower-right of the area covered by FORV-CONTOUR data.

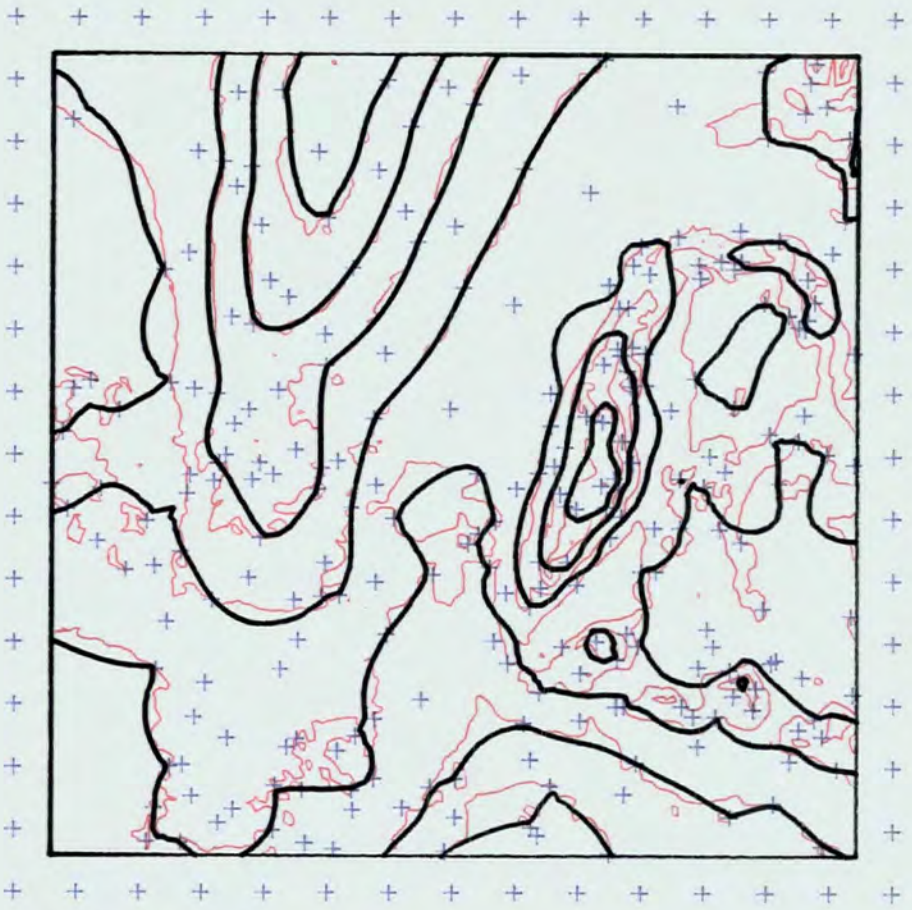
6.5 GLOBAL METHODS

6.5.1 GPCP-TREND

There are numerous global polynomial random-to-grid programs available. The 'TREND' routine of GPCP was chosen as it was considered the most flexible available; it allows interpolation of first- to tenth-order orthogonal polynomials from point data. In theory, higher, rather than lower, order surfaces should yield superior results (Walden, 1972, 44), and therefore seventh-order surfaces were generated. Higher-order polynomials could not be generated at NUMAC from the size of data samples used in this research due to implementation limitations.

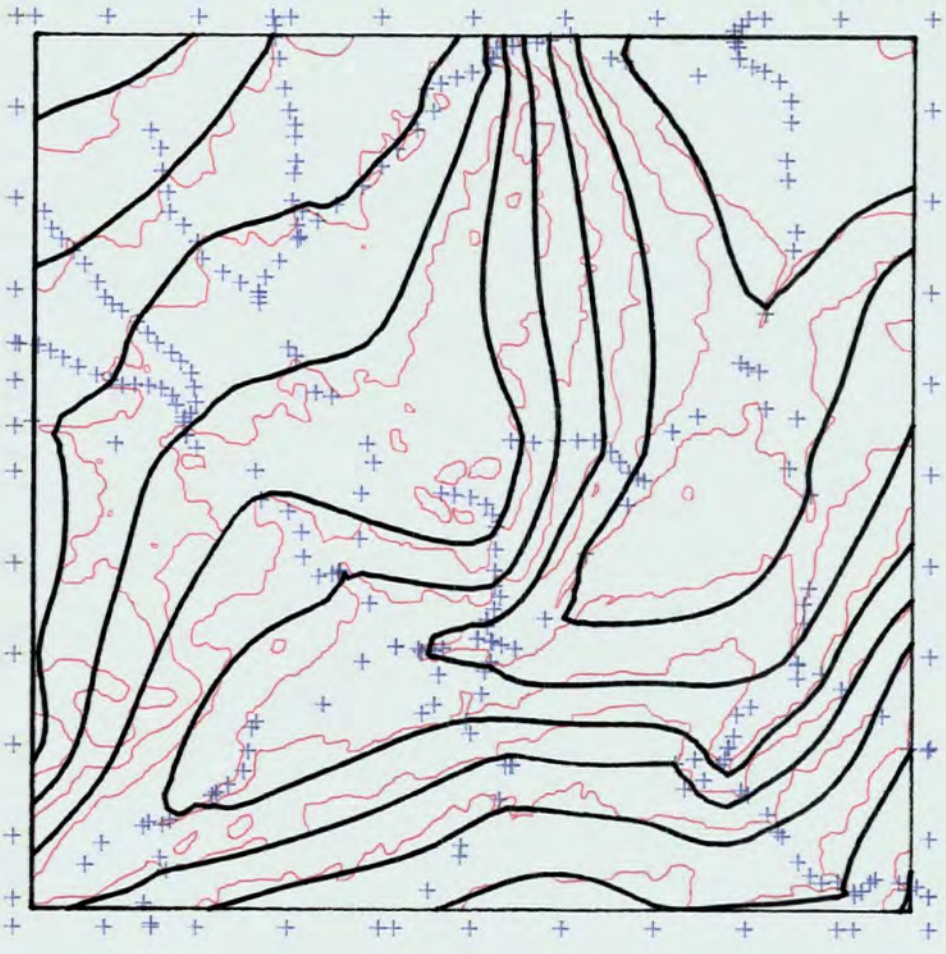
It is apparent from Table 6.17 that the thirty-six term polynomial cannot cope with the data sets so well as any of the previous methods, with the possible exception of MINCURV. Interpolation error is related to data set characteristics in addition to model characteristics. For example, all non-zero RBAD statistics only occur with INCH data and these are not

Figure 6.24



F-C MULTI

I-R MULTI



DS	%BAD	ORIGINAL		TRIMMED		TRIMMED ERROR				TRIMMED ABSOLUTE ERROR			CORRELATIONS			
		MIN	MAX	MIN	MAX	MEAN	MEDIAN	SKEWNESS	KURTOSIS	MEAN	MEDIAN	STANDARD	SURFACE	NEAR	NONN	GRADIENT
IB	1.54	-51.66	46.96	-32.97	32.88	-1.06	-1.36	0.29	3.02	10.38	8.30	13.14	0.951	0.104	-0.090	0.041
IC	0.04	-40.02	33.10	-26.88	26.36	-0.28	0.24	-0.15	2.68	8.39	7.29	10.43	0.970	0.015	-0.022	-0.018
IG	6.07	-62.22	61.89	-41.60	41.74	0.25	-0.36	-0.01	2.38	14.84	13.02	17.96	0.921	0.007	-0.028	0.151
IR	31.19	-107.61	113.75	-94.08	93.86	-5.06	-7.23	0.28	2.97	30.17	24.74	37.63	0.746	0.514	-0.523	0.001
IS	0.08	-43.97	41.47	-23.60	23.44	-0.37	-0.42	0.19	2.36	8.69	7.56	10.51	0.971	0.059	-0.071	0.026
LB	0.00	-8.64	6.77	-6.68	6.19	0.02	0.06	-0.24	3.22	1.60	1.31	2.04	0.996	-0.109	0.134	0.161
LC	0.00	-14.06	1.96	-8.31	1.76	-4.57	-4.76	0.36	2.57	4.59	4.77	4.98	0.996	-0.051	0.085	0.169
LG	0.00	-9.96	5.86	-3.77	3.76	-0.02	-0.04	0.01	2.65	1.19	1.19	1.48	0.998	0.011	0.025	0.202
LR	0.00	-16.07	5.98	-10.48	5.80	-2.11	-2.39	-0.56	2.69	2.96	2.06	4.00	0.988	0.652	-0.532	0.373
LS	0.00	-26.83	20.00	-18.72	18.05	-0.58	-0.62	-0.15	2.50	6.15	5.44	7.51	0.946	0.086	-0.158	-0.225
FB	0.00	-13.10	9.22	-7.55	7.60	1.09	1.06	0.19	3.22	2.14	1.69	2.88	0.865	0.259	-0.267	0.031
FC	0.00	-11.77	6.93	-5.27	5.29	1.21	1.39	-0.44	2.72	2.21	1.92	2.57	0.882	0.071	-0.022	-0.058
FG	0.00	-14.76	5.48	-4.93	4.73	0.24	0.53	-0.28	2.14	1.98	1.86	2.34	0.888	-0.001	0.035	-0.006
FS	0.00	-12.43	8.85	-7.00	7.02	1.15	1.31	-0.17	3.00	2.29	1.97	2.82	0.868	0.118	-0.167	0.006

Table 6.17a Primary results of GPCP-TREND

DS	IN-NEAR			IN-NONN			IN-GRAD			OUT-NEAR			OUT-NONN			OUT-GRAD		
	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN	MIN	MAX	MEAN
IB	0.95	174.83	50.61	0	14	3.80	0.12	83.55	19.63	6.36	168.94	70.47	0	11	2.39	1.86	52.75	14.68
IC	0.83	182.47	44.19	0	10	3.32	0.17	83.55	19.46	3.02	108.16	42.24	0	8	3.19	0.12	65.32	17.89
IG	0.00	42.42	32.38	4	6	4.19	0.12	83.55	19.41	0.00	42.42	33.50	4	6	4.22	2.11	54.61	18.66
IR	0.91	235.28	68.94	0	20	3.47	0.17	83.55	19.08	30.00	306.41	198.52	0	1	0.07	0.12	52.38	25.19
IS	0.31	150.02	47.02	0	12	3.31	0.12	83.55	19.08	1.88	100.97	42.19	0	9	3.59	1.11	66.82	25.16
LB	0.31	165.00	40.10	0	23	3.11	0.00	41.50	10.38	4.00	21.21	11.04	0	18	9.09	4.68	27.72	15.69
LC	0.00	133.99	28.97	0	11	3.18	0.00	41.50	10.14	1.06	106.60	18.14	0	10	4.81	1.67	39.77	15.46
LG	0.00	42.42	19.80	1	6	3.32	0.00	40.00	10.15	0.00	42.42	19.99	1	5	3.32	2.68	41.50	15.06
LS	0.31	226.99	51.64	0	19	3.42	0.00	41.50	10.02	37.10	226.99	174.08	0	3	0.02	4.14	33.14	17.69
LS	0.21	112.92	27.05	0	11	3.11	0.00	41.50	10.71	2.46	88.13	39.91	0	6	1.61	0.35	17.11	4.64
FB	0.21	114.44	26.94	0	18	4.81	0.00	84.47	8.64	1.31	129.47	75.00	0	12	1.06	0.13	56.60	7.80
FC	0.40	70.00	21.46	0	16	4.60	0.00	84.47	8.54	1.13	42.47	20.41	1	17	5.43	0.28	56.60	9.93
FG	0.00	41.22	16.39	2	7	4.75	0.00	74.47	8.11	0.00	31.61	16.58	2	7	4.71	0.40	84.47	17.85
FS	0.00	80.61	23.72	0	17	4.79	0.00	84.47	8.67	3.26	84.33	41.91	0	10	1.82	0.12	56.60	7.43

Table 6.17b Secondary results of GPCP-TREND

solely related to one data characteristic - GRID data with the highest NNS has the second highest %BAD value, whereas the low NNS RIVER data set has a staggering 31.2% nodes not satisfactorily interpolated.

The ORIGINAL and TRIMMED MIN/MAX statistics underline this variability. In addition to having the second highest mean magnitudes (ORIGINAL - 28.66, TRIMMED - 20.38), there is a high within-model variance with a consistent negative bias. Unlike most other methods, LIAN-RIVER data produce surprisingly good results, but the equivalent INCH data set produce poor results. The exact opposite effect was produced by the use of SCATTER and GRID data.

The other precision descriptor, KURTOSIS, behaves uniquely. With such a large TRIMMED ERROR range, it would be expected that KURTOSIS would be increased from the previous methods, however this is the only interpolation method which has a mean KURTOSIS value less than 3.0 (2.73). This suggests a more even distribution of error over the error range than for all the previous methods.

The MEAN and MEDIAN statistics support the previous theory concerning a strong negative bias, although this is highly variable within each model and between models. Use of INCH and LIAN data produces several results with a strong negative bias component; however all results based on FORV data are strongly positively biassed.

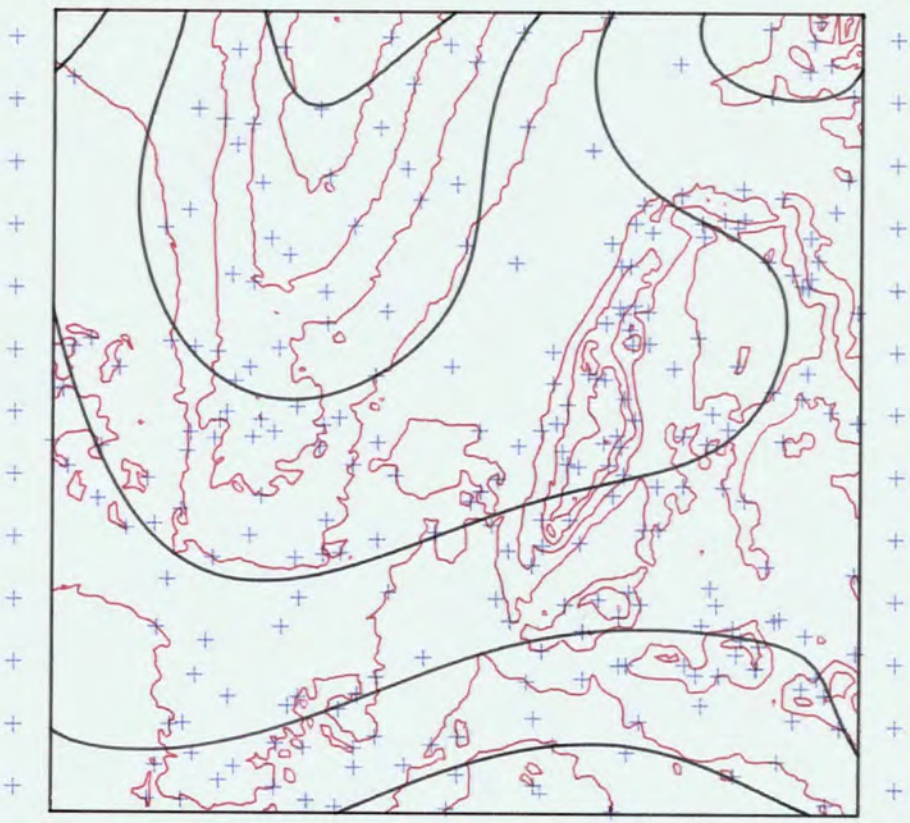
All the accuracy statistics are similarly variable. There is strong consistency within the results from the FORV data

sets, although the LIAN and INCH models produce a high within-model variance. In addition, mean ABS MEAN (6.98) and mean STANDARD ERROR (8.60) are the second worst encountered in this study. Importantly, there is no significant rank correlation between NNS and any of the accuracy descriptors (STANDARD ERROR - 0.5), although a stronger correlation exists with vector dispersion analysis and surface area analysis.

There is no doubt that the error variability encountered within trend surface fitting is a significant problem, and prevents a good appreciation of error sources. It is a problem similar to that encountered within the patchwise BRAILE method where it was suggested that the error variability increased with patch size. Within GPCP-TREND, the patch has been increased to cover the whole model area, and the polynomial order has been similarly increased. While this leads to the creation of few extreme errors, it is impossible for the polynomial to satisfactorily model complex topography, resulting in few low errors and many moderate errors - hence the sub-normal KURTOSIS statistics.

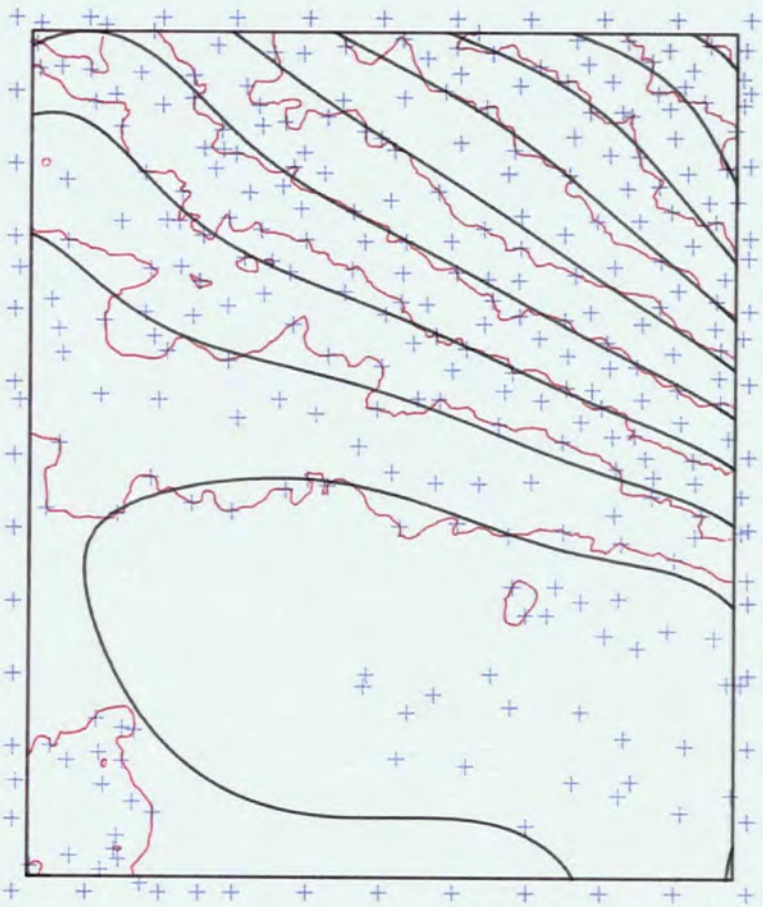
The full effect of using a trend surface is illustrated in Figure 6.25. While no major depressions or peaks exist as in GINOSURF and GHOT, the resultant isarithms have very little similarity with reality where irregular surfaces exist (eg. FORV-CONTOUR). Not surprisingly, some amount of similarity may be achieved in modelling a more regular surface (eg. LIAN-CONTOUR).

Figure 6.25



F-C GP7T

L-C GP7T



6.6 CONCLUSION

Throughout the previous discussion, the emphasis has been on discovering the statistically optimum method of random-to-grid interpolation for specific data sets. However, only in a few circumstances - for example in mineral exploration - can other processing considerations of these methods be disregarded. Additionally, the importance of this research is not purely in establishing the relative accuracies of specific situations, but lies in the possible prediction of the optimum method of interpolation for a given data set.

In concluding this chapter, four primary aspects of the interpolation process will be considered in more depth. The previous statistical results will be summarised and briefly examined for any trends. Processing limitations will be investigated, particularly with reference to the memory requirements of each technique, the complexity of the form of the input, and the requirements of processing time and cost. An attempt will be made to associate the statistical results with the surface and data characteristics.

These four primary aspects form a basis from which the general user may decide on the optimum method for his situation.

6.6.1 Summary Of The Statistical Results

The statistical results have already been analysed in considerable detail. This section attempts purely to summarise

the analysis, and is based on a consideration of the major statistical descriptors discussed earlier (consistency, bias, precision and accuracy). Table 6.18 provides a relative summary of all the methods for each of the descriptors. The methods have been graded in the range 1 to 5, the distribution of gradings being normalised: high values indicate good performance. An overall performance total has been subsequently derived, assuming each factor has the same weight. Surprisingly there is a remarkable grade consistency for each method.

The consistency measures (%BAD and TRIMMED MIN/MAX) are most associated with interpolation range and search method. Consistency is generally excellent where patchwise and hybrid methods of interpolation are utilised. The improved search ranges, in association with a more rigorous interpolation method, ensure that error is better constrained in data voids (limiting extreme errors) than in other methods. Where simple pointwise interpolation is used, consistency is directly related to the search method - improving with directionality and search distance (SURFACE II GRAPHICS methods). As more complex pointwise functions are utilised (GINOSURF, GHOT, KRIGE and MINCURV), the interpolation method is more susceptible to extremes of data distribution and surface morphology.

The bias descriptors (MEAN and MEDIAN) have been dramatically affected by INCH-RIVER data and illustrate a polarisation with model type - results from FORV data are predominately negatively biassed, those from LIAN data

METHOD	CONSISTENCY	BIAS	PRECISION	ACCURACY	MORPH. TRUENESS	PERFORMANCE TOTAL	PERFORMANCE ORDER
SII-N	2	4	3	3	3	15	8
SII-Q	3	4	4	4	4	19	4=
SII-O	3	5	3	3	4	18	6=
SII-V	4	5	3	3	4	19	4=
GPCP	3	3	3	3	2	14	9
SYMAP	4	2	2	2	3	13	10=
GINOSURF	3	4	4	4	3	18	6=
GHOT	2	1	2	2	2	9	12
KRIGE	2	3	5	5	5	20	2=
MINCURV	1	1	1	1	1	5	14
SACM	5	2	4	4	5	20	2=
BRAILE	4	3	2	2	2	13	10=
MULTI	5	3	5	5	3	21	1
GPCP-7T	1	2	1	1	1	6	13

Table 6.18 Summary of the statistical results of the various interpolation methods.

predominately positively biassed, and those from INCH data being variable. However, if a mean is established for each method, it is apparent that the simpler functions are most successful. This is not surprising since, once extreme errors have been removed by using trimmed distributions, the simpler data averaging methods (SURFACE II GRAPHICS methods) should be most representative of the surface, assuming a good distribution of data points is used (note the improvement with the OCTANT and VRADIUS methods). More complex functions are much more susceptible to extremes of data distribution.

The precision (TRIMMED MIN/MAX and KURTOSIS) and accuracy (STANDARD ERROR, ABS MEAN, ABS MEDIAN, SURFACE CORRELATION) statistics may be considered together since, as a result of the generally low magnitude systematic error, they are similar. Once again, the simpler functions have performed moderately well. The elimination of the trimmed extreme errors has improved the position of the more specialised functions, for example GINOSURF and KRIGE: however, the methods which generate a large amount of random noise (GPCP-TREND, SYMAP and BRAILE) have remained relatively constant while the other methods have improved. GPCP-TREND, SYMAP and BRAILE continue to generate interpolation noise in areas where a good interpolation is possible, though for very different reasons.

The morphological trueness is more difficult to measure quantitatively without performing a complex and detailed examination - see the evaluation of the multiquadric interpolated grids in Chapter 7. In general, and excluding the methods which generate strong, distinctive patterns (e.g. the

QUADRATIC EFFECT, the LINEAR BLOCK ISARITHM EFFECT and the PATCHWISE EFFECT), superior morphological trueness occurs with the more localised pointwise methods and those more complex functions which pay more attention to localised trends - e.g. KRIGE and SACM. The more global functions, which utilise a wide search to determine which data points to use in the interpolation, tend to over-smooth the resultant grid due to the wide sphere of influence of each data point. This need not necessarily be a disadvantage, and may be an advantage in certain cases (see 2.2).

The overall performance total, while perhaps not emphasising the distinct superiority of the better interpolators due to the grading process used, provides a good 'rule-of-thumb' to the optimum interpolation methods. Clearly MINCURV, GPCP-TREND and GHOT must be discounted as optimum interpolators for topographic data. The poor control over the interpolation function and data requirements for GHOT, and the over-smoothing of the MINCURV and GPCP-TREND methods have been magnified by the variable data point distribution and the surface roughness encountered. The specialised methods which pay particular importance to honouring the data points are most successful (KRIGE, SACM and MULTI). GINOSURF and MULTI also underline the success of using quadratic functions, in which the process may be locally constrained by using a wide distribution of data per individual interpolation. Finally, the consistent though moderate success of using simple data averaging is shown when a balanced search procedure is utilised.

6.6.2 Processing Considerations

While the major consideration in comparing interpolation methods is on the quality of the resultant product, very often the actual logistics of running and accessing the programs must be considered. A major limitation is very often the memory requirements of a specific program, since small non-virtual systems may not be able to cope with large task sizes successfully. The processing time may also be a severe limitation if the process is to be performed many times, although if the interpolation is a 'one-off' it should be of less importance. Additional requirements, perhaps involving pre-processing, must be considered in conjunction with both these aspects. These considerations may be summarised in terms of cost. While it is unlikely that this would be a controlling factor, it may be of importance if use of a bureau facility was being contemplated.

Summary statistics for running each random-to-grid interpolation method have been derived and are presented in Tables 6.19 and 6.20. The statistics represent those for the FORV-SCATTER data (approximately 300 points), although these are consistent with most of the other data sets. The statistics (Table 6.19), with the exclusion of SACM, SURFACE II GRAPHICS and KRIGE, represent those quoted at NUMAC for the IBM 370/168. Approximate comparable estimates have been established for the other packages which were run on external sites. These statistics do not include costs in respect of data I/O, printer and plotter usage, file storage and elapsed time. However, the costs of compiling private programs and

Package	Main Run		Associated		Total	
	Memory	Time	Memory	Time	Memory	Time
SURFACE II	9.8	9.5			9.8	9.5
GPCP	16.37	10.55	2.12	4.14	18.49	14.69
SYMAP	10.05	8.51	1.83	4.09	11.88	12.60
GINOSURF	15.90	13.48	0.68	0.84	16.58	14.32
GHOT	24.83	34.96	3.32	2.19	28.15	37.15
KRIGE	280.00	255.00			280.00	255.00
MINCURV	3.35	4.86	18.63	9.41	21.98	14.27
SACM	14.30	11.50			14.30	11.50
BRAILE	5.20	5.22	6.67	3.90	11.87	9.12
MULTI	278.59	76.77	1.30	1.17	279.89	77.94
GPCP-TREND	19.03	12.33	2.12	4.14	21.15	16.47

Table 6.19 CPU time (seconds) and virtual memory (page/minute) requirements for random-to-grid programs.

Package	Main Run		Associated		Total Cost
	Memory	Time	Memory	Time	
SIIG	2.8	256.5			259
GPCP	4.2	284.9	1.4	112.0	402
SYMAP	3.2	229.8	1.2	110.4	345
GINOSURF	3.2	364.0	2.2	22.7	392
GHOT	1.9	943.9	4.1	59.1	1009
KRIGE	3.0	6885.0			6888
MINCURV	1.9	131.2	5.3	254.1	393
SACM	3.4	310.5			314
BRAILE	2.7	140.9	4.6	105.0	254
MULTI	9.8	2072.8	3.0	31.6	2117
GPCP-TREND	4.2	332.9	1.4	111.8	450

Table 6.20 Run costs (in pence) for random-to-grid interpolation.

data I/O reformating (for GPCP and SYMAP) are included within associated costs. The run statistics have been converted into costs (in pence) as based on NUMAC commercial rates (see Table 6.20).

In general, the costs (Table 6.20) are highly variable. The majority of packages cost in the region of 350p per run. However, by far the largest charges refer to GHOT, MULTI and KRIGE. From the previous discussion, it is apparent that neither GHOT nor KRIGE justifies the expense from their performances; however MULTI cannot be disregarded so easily. Clearly, aspects of the data and the required resultant grid accuracies must be carefully examined before a decision to use either MULTI or SACM (12% of the cost of MULTI) is made.

If the costs are examined and broken down into their constituent parts, additional considerations must be made. BRAILE and MINCURV incur substantial costs of compilation of FORTRAN source, which - if the program were used regularly - would not be required, since the compiled program would be stored. If these costs are removed, the total costs would be approximately 150p - a very attractive proposition if a suitable data set is being used. Alternatively, some programs require excessive amounts of CPU memory, which may prove restrictive in smaller computers. Unfortunately, the maximum capacity used could not be derived at run-time. However, since 1 MTS page is equivalent to 4096 bytes, approximate virtual memory sizes may be established - MULTI (1.5mbytes), GPCP and GINOSURF (400kbytes), SACM (300kbytes), BRAILE (240kbytes) and GHOT (170kbytes).

On the basis of these results, serious consideration must be given by potential users as to whether the required accuracy or data characteristics warrant use of a certain package and to whether certain data sizes or CPU restrictions will permit it to be run in the local computer environment. For example, though BRAILE and MINCURV are only attractive where well-distributed data exist and SACM and SURFACE II GRAPHICS are particularly attractive where the data are poorly distributed, reliance on MULTI may be necessary in cases of extreme data distribution problems.

6.6.3 Surface Characteristic Considerations

For the purposes of being able to predict the success of the various random-to-grid interpolation methods for any data sample, summary statistics for each interpolation method were correlated with summary statistics from the surface characteristic examination (see Table 6.21-6.23). Three summary interpolation statistics were examined to consider the possible prediction of:

a. interpolation consistency (%BAD);

b. interpolation systematic error (MEAN ERROR);

c. interpolation accuracy (STANDARD ERROR).

The surface characteristics examined were:

PROG	AUTO GRID	AUTO PT	SA GRID MEDIAN	SA PT MEAN	VDA GRID V	ALT GRID MEAN	GRAD GRID S.D.	PROF GRID S.D.	PLAN GRID S.D	NNS PT
SII-N	0.30	0.12	-0.25	-0.19	-0.14	-0.08	-0.15	-0.14	-0.39	-0.52
SII-Q	-0.05	0.02	0.25	0.36	0.22	0.22	0.24	-0.20	-0.09	-0.38
SII-O	0.26	0.21	0.05	0.15	0.11	0.17	0.10	-0.34	-0.50	-0.57
SII-V	0.15	0.24	0.38	0.51	0.39	0.40	0.37	-0.47	-0.38	-0.42
GPCP	-0.03	0.05	0.30	0.41	0.27	0.27	0.28	-0.25	-0.11	-0.37
SYMAP	0.14	0.20	0.29	0.42	0.30	0.32	0.29	-0.40	-0.39	-0.45
GINOSURF	0.09	0.02	0.04	0.16	0.06	0.10	0.08	-0.23	-0.18	-0.58
GHOT	0.30	0.09	-0.15	-0.03	-0.06	0.01	-0.04	-0.25	-0.31	-0.60
KRIGE	0.51	0.41	0.14	0.22	0.24	0.31	0.24	-0.52	-0.71	-0.58
MINCURV	0.46	0.23	0.14	0.34	0.21	0.27	0.23	-0.52	-0.29	-0.47
SACM	0.10	0.18	0.37	0.48	0.35	0.36	0.35	-0.41	-0.29	-0.35
BRAILE	0.04	0.10	0.23	0.34	0.22	0.24	0.23	-0.27	-0.24	-0.49
MULTI	0.10	0.18	0.37	0.48	0.35	0.36	0.35	-0.41	-0.29	-0.35
GPCP-7	0.17	0.29	0.43	0.44	0.42	0.44	0.43	-0.48	-0.33	-0.30

Table 6.21 Interpolation method/surface characteristic cross-correlations (%BAD)

PROG	AUTO GRID	AUTO PT	SA GRID MEDIAN	SA PT MEAN	VDA GRID V	ALT GRID MEAN	GRAD GRID S.D.	PROF GRID S.D.	PLAN GRID S.D	NNS PT
SII-N	-0.24	-0.25	-0.27	-0.38	-0.30	0.70	-0.31	0.46	0.47	0.34
SII-Q	-0.19	-0.24	-0.33	-0.44	-0.34	0.69	-0.34	0.45	0.41	0.30
SII-O	-0.19	-0.24	-0.33	-0.44	-0.34	0.65	-0.34	0.45	0.42	0.30
SII-V	-0.20	-0.24	-0.32	-0.42	-0.33	0.71	-0.33	0.46	0.42	0.30
GPCP	-0.30	-0.27	-0.30	-0.39	-0.34	0.63	-0.34	0.56	0.49	0.46
SYMAP	-0.36	-0.43	-0.56	-0.60	-0.58	0.79	-0.58	0.66	0.52	0.35
GINOSURF	-0.19	-0.16	-0.27	-0.35	-0.28	0.60	-0.29	0.47	0.37	0.37
GHOT	-0.27	-0.22	-0.23	-0.35	-0.27	0.43	-0.27	0.51	0.47	0.46
KRIGE	0.29	0.47	0.74	0.76	0.72	0.56	0.71	-0.66	-0.43	-0.45
MINCURV	-0.49	-0.47	-0.46	-0.54	-0.51	0.22	-0.51	0.70	0.58	0.55
SACM	-0.19	-0.20	-0.31	-0.40	-0.32	0.59	-0.32	0.47	0.38	0.34
BRAILE	-0.21	-0.26	-0.39	-0.47	-0.39	0.75	-0.39	0.49	0.43	0.34
MULTI	-0.22	-0.23	-0.32	-0.42	-0.33	0.68	-0.34	0.50	0.42	0.38
GPCP-7	-0.59	-0.56	-0.44	-0.51	-0.54	0.79	-0.51	0.70	0.80	0.50

Table 6.22 Interpolation method/surface characteristic cross-correlations (MEAN ERROR).

PROG	AUTO GRID	AUTO PT	SA GRID MEDIAN	SA PT MEAN	VDA GRID V	ALT GRID MEAN	GRAD GRID S.D.	PROF GRID S.D.	PLAN GRID S.D	NNS PT
SII-N	0.28	0.51	0.84	0.85	0.81	0.29	0.79	-0.68	-0.37	-0.87
SII-Q	0.29	0.52	0.80	0.82	0.79	0.26	0.76	-0.68	-0.43	-0.84
SII-O	0.24	0.49	0.84	0.85	0.81	0.22	0.78	-0.66	-0.37	-0.87
SII-V	0.32	0.53	0.78	0.81	0.77	0.27	0.75	-0.69	-0.47	-0.85
GPCP	0.28	0.45	0.73	0.78	0.72	0.28	0.70	-0.65	-0.40	-0.85
SYMAP	0.09	0.39	0.75	0.73	0.70	0.19	0.67	-0.54	-0.29	-0.66
GINOSURF	0.33	0.49	0.71	0.77	0.71	0.43	0.69	-0.68	-0.47	-0.78
GHOT	0.33	0.35	0.68	0.62	0.68	0.15	0.67	-0.67	-0.44	-0.69
KRIGE	0.24	0.44	0.72	0.74	0.70	0.25	0.69	-0.63	-0.39	-0.70
MINCURV	0.61	0.68	0.75	0.82	0.79	0.41	0.79	-0.82	-0.62	-0.81
SACM	0.23	0.44	0.76	0.79	0.73	0.20	0.72	-0.62	-0.36	-0.73
BRAILE	0.28	0.56	0.89	0.85	0.86	0.17	0.84	-0.70	-0.39	-0.74
MULTI	0.21	0.41	0.71	0.73	0.68	0.23	0.67	-0.60	-0.36	-0.70
GPCP-7	0.34	0.65	0.91	0.77	0.88	0.20	0.87	-0.72	-0.42	-0.46

Table 6.23 Interpolation method/surface characteristic cross-correlations (STANDARD ERROR).

a. two-dimensional autocorrelation (the vector statistic);

b. areal autocorrelation (the curve decay distance when the function is 0.707);

c. surface area analysis (grid median);

d. surface area analysis (point median);

e. vector dispersion analysis (the 'V' statistic);

f. pointwise polynomial (mean altitude);

g. pointwise polynomial (standard deviation of gradient);

h. pointwise polynomial (standard deviation of profile convexity);

i. pointwise polynomial (standard deviation of plan

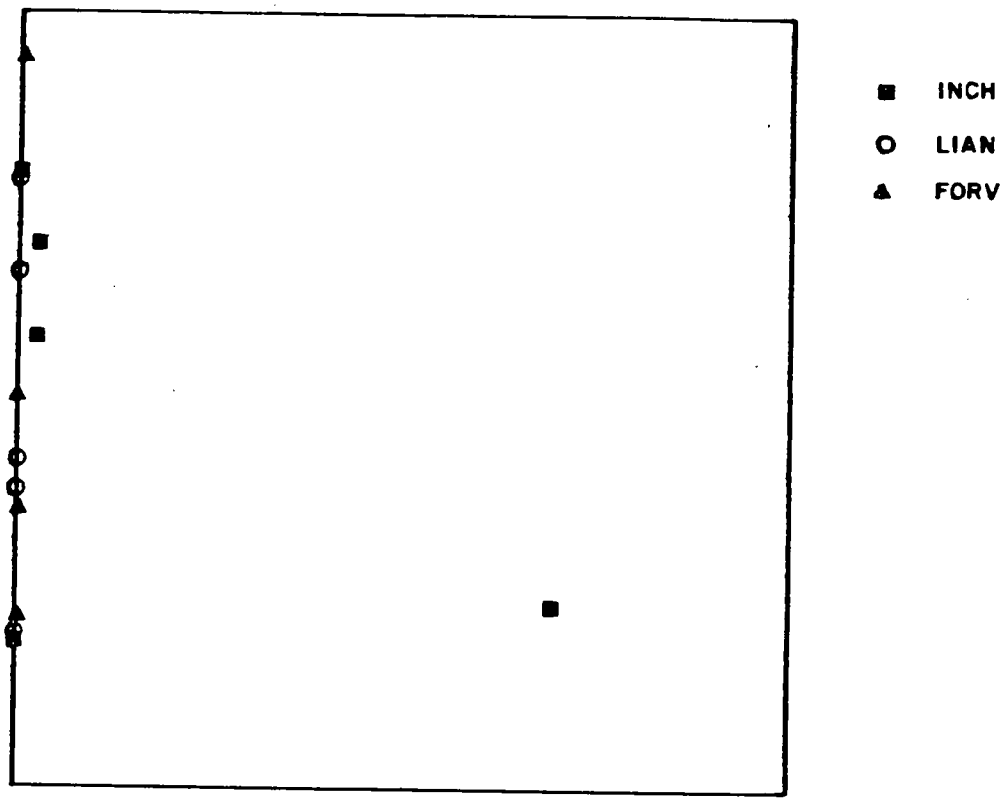
convexity).

Though many other surface characteristic statistics were examined, the majority were disregarded giving this important, yet representative, sample.

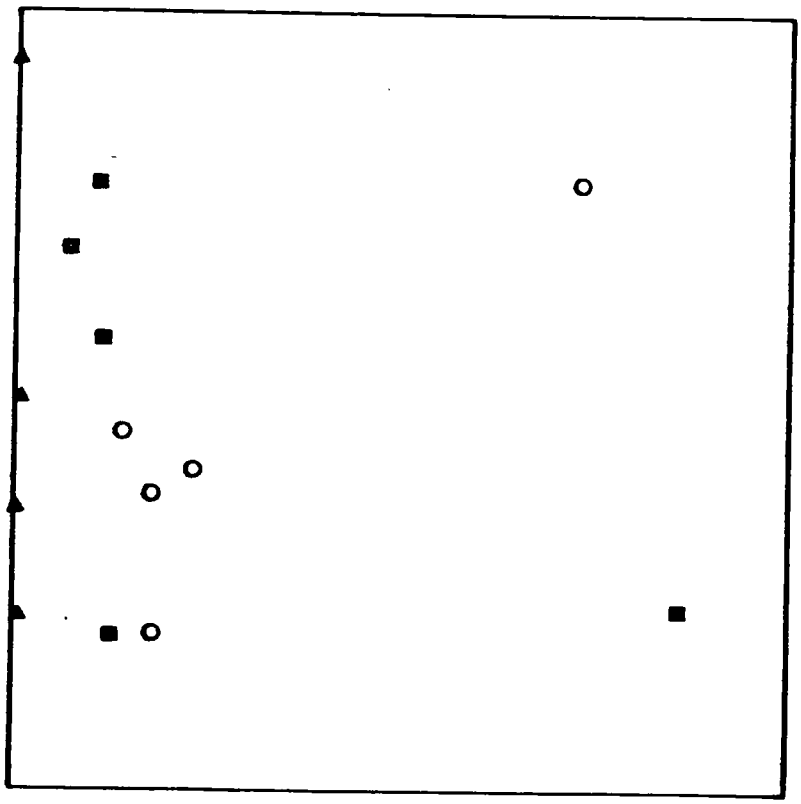
From the resultant Pearson's product moment coefficients of correlation, various conclusions may be made which are supported by a few randomly selected scattergrams (Figure 6.26-28).

a. There are no important correlations between consistency of interpolation and the surface characteristics. While this may be a feature of the association between the variables (perhaps variable association for each model surface - see Figure 6.26b), it is more likely caused by the high number of zero-case %BAD values - any distribution truncated at a fixed value (see Figure 6.26a) is likely to cause problems in evaluating a high cross-correlation with another distribution.

b. There are no significant correlations between systematic error and the surface characteristics. Interesting correlations do occur with the areal-autocorrelation and more importantly the convexity-based correlations. However, the most important correlations are with altitude (Figure 6.27a shows surprisingly good correlation strength for INCH and LIAN marred by a much poorer FORV correlation). As mentioned previously, higher correlations occur between altitude

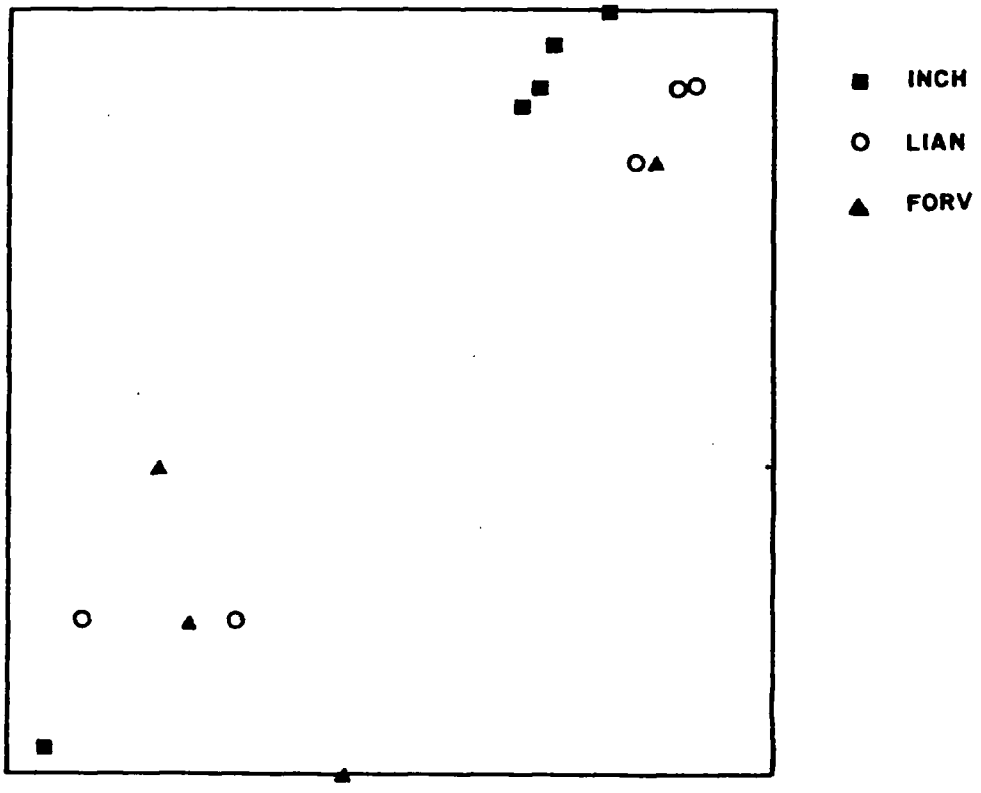


a) Two dimensional autocorrelation against GPCP-%BAD

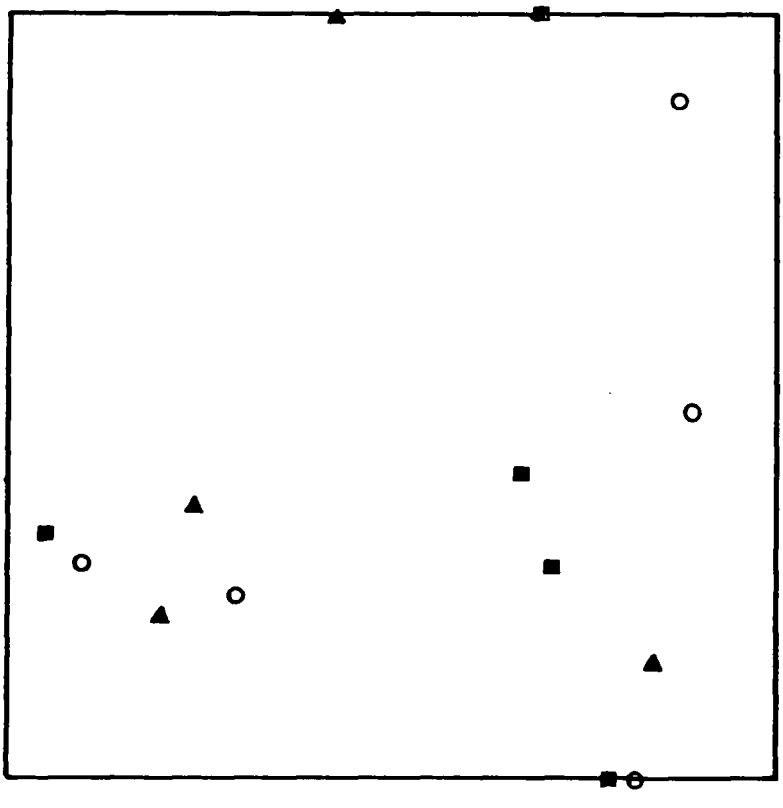


b) Two-dimensional autocorrelation against KRIGE-%BAD

Table 6.26 Scattergrams of consistency vs. characteristics.

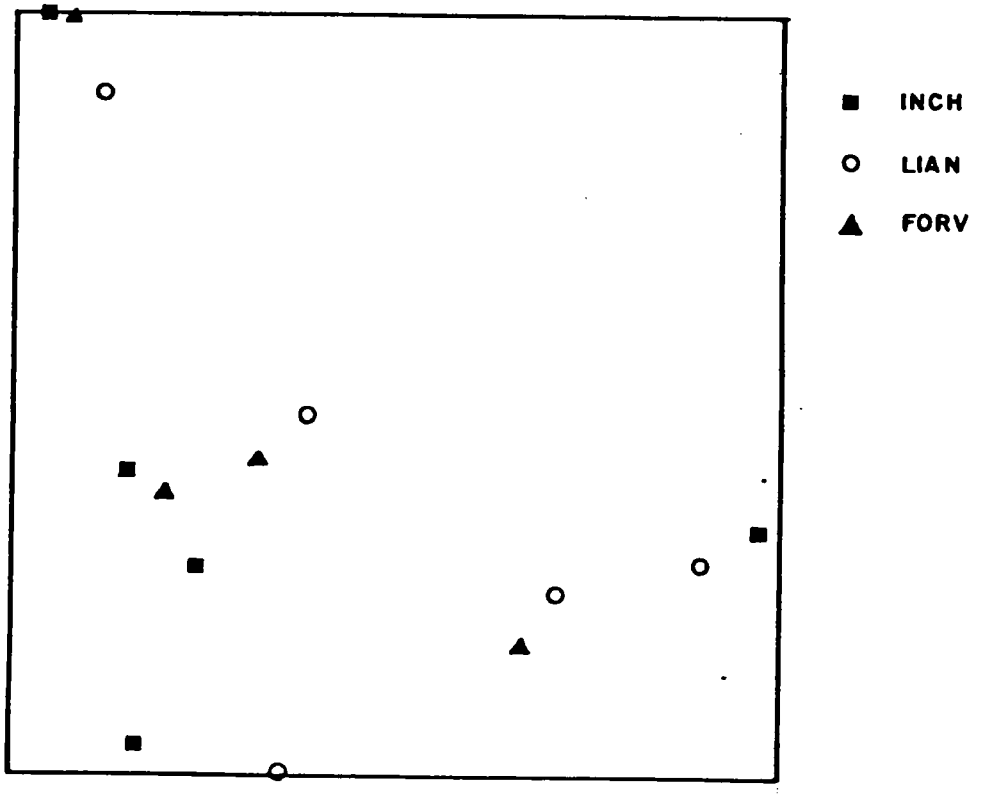


a) Mean attitude against GPCP-MEAN

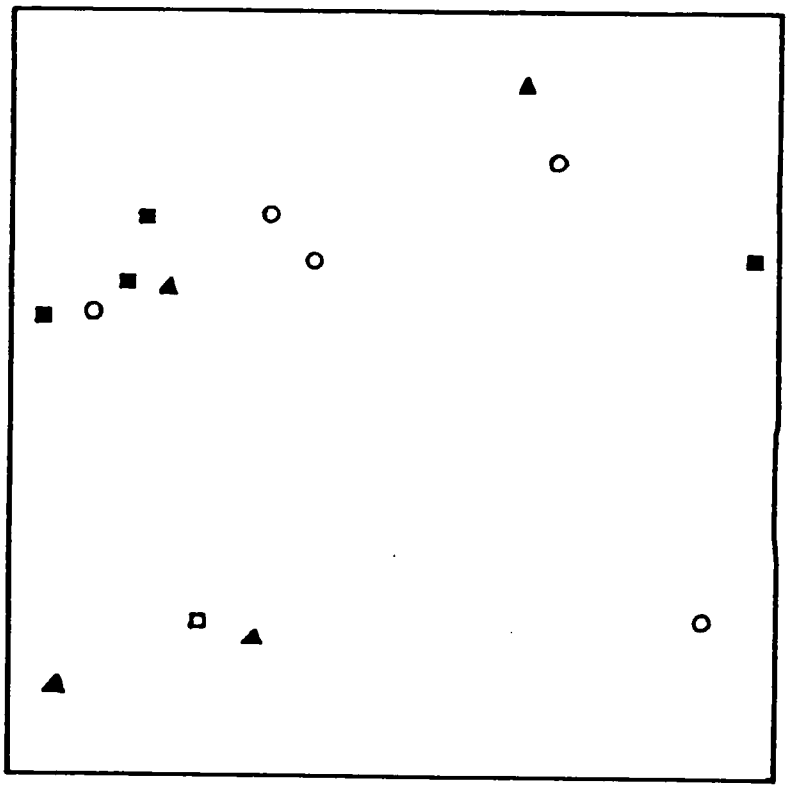


b) NNS against GPCP-MEAN

Table 6.27 Scattergrams of systematic error versus characteristics



a) NNS against GPCP - STANDARD ERROR



b) Two- dimensional autocorrelation against GPCP-STANDARD ERROR

Figure 6.28 Scattergrams of accuracy vs. characteristics.

and the interpolation methods which produce most random noise - SYMAP, BRAILE and GPCP-TREND.

c. There are several significant correlations between accuracy and the surface characteristics - most noticeably with the surface area analyses, vector dispersion analysis and gradient. Importantly, these are all associated with the first derivative of altitude, and it is therefore not surprising that they are most highly correlated with the interpolation methods which utilise local linear slopes. This suggests that interpolation accuracy is directly related to both the magnitude and the distribution of surface gradient. In contrast, the surface characteristics based on the second derivative of altitude - those based on the transition of local surface gradients (the autocorrelation and convexity characteristics) - perform less well as predictors (see Figure 6.28b - several outliers for each model, especially FORV, have combined with high cross-model variance to give a poor global correlation). It must be concluded that slope distribution and magnitude are more important than slope transition for predicting interpolation performances.

Another approach is to study the reactions of the individual model surfaces to particular interpolation methods by considering the mean consistency, precision, accuracy and bias statistics for each surface. Individual model means were generated from each of the relevant data sets for each of the interpolation methods for consistency (ORIGINAL MIN/MAX),

precision (KURTOSIS), accuracy (STANDARD ERROR) and bias (MEAN). For bias, the RIVER values were omitted for INCH and LIAN since they are inherently strongly negatively biassed. The results are summarised in Table 6.24.

Several general conclusions may be drawn from Table 6.24 in relation to the general success of random-to-grid interpolation.

a. Flat and/or smooth data surfaces (as described by surface area analysis, vector dispersion analysis and pointwise polynomial-gradient) produce higher accuracy and consistency statistics than do rougher data surfaces.

b. Surfaces with considerable micro-relief (eg. represented by the FORV data) produce higher precision statistics. Micro-relief surfaces (as delineated by areal autocorrelation) result in generally lower error due to the global smoothness of the surface, although the local nuances may result in a few extremes which tend to increase KURTOSIS.

c. While the smooth surfaces, which generally have low associated interpolation error, exhibit consistent strong positive (FORV) and negative (LIAN) bias, the rough macro-featured surface (INCH) exhibits more variable bias, and therefore may be considered a superior indicator of the biassing effect of individual interpolation methods.

METHOD	INCH				LIAN				FORV			
	accuracy	bias	consistency	precision	accuracy	bias	consistency	precision	accuracy	bias	consistency	precision
SII-N	5.63	0.45	32.52[3]	3.32	1.58	0.10	9.33[4]	3.37	1.43	0.54	8.35[2]	3.82
SII-Q	5.33	0.39	28.77	3.24	1.72	0.18	9.00[1]	3.43	1.56	0.61	8.23[1]	3.52
SII-O	5.42	0.36	29.48	3.21	0.63	0.30	7.95[3]	3.23	1.85	0.76	8.43[1]	3.39
SII-V	5.59	0.44	14.97[2]	3.33	1.87	0.14	10.31	3.52	1.46	0.56	7.93	3.64
SYMAP	6.60	-1.04	37.26	3.32	1.86	-0.38	12.91	3.30	3.19	0.64	14.56	3.22
GPCP	6.02	0.10	44.65	3.72	1.73	-0.48	11.90	3.52	1.34	0.11	12.54	4.56
GINOSURF	4.82	0.26	40.01	3.51	1.70	-0.26	56.73[1]	3.71	1.07	0.12	16.49	4.62
GHOT	7.17	0.03	53.46[2]	3.59	3.50	-0.68	158.32[2]	6.84	1.55	0.22	67.88	6.80
KRIGE	4.29	0.10	21.48[10]	3.23	1.31	-0.34	7.03[10]	3.62	0.96	0.19	7.40[4]	4.06
MINCURV	20.79	-1.48	215.77	8.87	40.09	9.53	151.49	6.86	1.10	0.32	7.87	3.76
SACM	4.59	0.24	24.81	3.45	1.49	-0.24	8.00	4.67	1.11	0.28	8.47	4.55
BRAILE	7.36	-0.10	47.53[1]	3.20	1.56	0.17	12.83	2.83	1.60	0.45	10.84	3.23
MULTI	4.35	0.09	20.29	3.04	1.16	-0.29	6.16	3.05	0.99	0.21	6.96	4.17
GPCP-TREND	17.94	-0.37	60.27	2.69	4.01	-1.29	11.62	2.73	2.66	0.93	10.32	2.77

Table 6.24 Summary of mean, bias, accuracy and precision for each model and for each interpolation method.

(Superscripts denote the number of extreme values omitted from the meaning due to resultant extreme statistical distortion).

In addition, more specific conclusions may be drawn.

a. All surfaces are consistently interpolated within SURFACE II GRAPHICS, irrespective of the search method used, unless too much directionality is applied.

b. Where an octant search is applied (OCTANT, KRIGE), FORV data are more susceptible to a decrease in accuracy or consistency. This is perhaps a direct result of the non-isotropic nature of the surface, which has a rapid decay in autocorrelation in the E-W direction, and a slower decay in the N-S direction.

c. GPCP is an accurate interpolator for planar surfaces (LIAN data). Where the surface displays either micro- or macro-roughness, the interpolation is generally good, although sharp morphological features may cause serious problems.

d. The quadratic pointwise methods (GINOSURF and GHOT) are not suited to smooth linear surfaces (those with high two-dimensional autocorrelation - e.g. LIAN). They perform better in micro- and macro-rough surfaces (FORV and LIAN) where surface irregularities suit the parabolic nature of the function.

e. KRIGE, which is based on considering correlograms generated from the data, is more suited to the surfaces with slow autocorrelation decay (as in INCH and LIAN data) and unsuited to those with rapid decay in autocorrelation (e.g. FORV data).

f. Similarly, the pseudo-splining MINCURV method is a good interpolator of smooth (as delineated by surface area analysis, vector dispersion analysis and pointwise polynomial-gradient) surfaces (e.g. FORV data) and is certainly not suited to use with rough surfaces.

g. In addition, the methods which utilise extensive patchwise and global low-order polynomials, and thus produce inherently smooth interpolated surfaces, are not suited to rough surfaces but are more suited to smooth surfaces.

h. Where superior interpolators are used (e.g. SACM and MULTI), the more variable surfaces improve most and the less variable surfaces, which have less scope for improvement, are improved to a lesser extent.

6.6.4 Data Considerations

As for surface characteristic considerations, the

consistency, systematic error and accuracy summary statistics for each interpolation method were correlated with the NNS for each data set (see Table 6.21-6.23 and Figures 6.27 and 6.28) in an attempt to predict the success of the various random-to-grid interpolation methods for any data sample. From the results, various conclusions may be made.

a. No significant correlations exist between %BAD and NNS, although there is a relatively consistent -0.45 (mean) correlation which is the most important %BAD correlation encountered. As with surface characteristics, this may be a direct result of the high number of zero case %BAD values.

b. No important correlation exists between MEAN ERROR and NNS. Figure 6.27 shows the dispersed nature of the NNS/GPCP-MEAN scattergram as against the stronger correlation and more lineated nature of the mean altitude/GPCP-MEAN scattergram (discussed in the previous section as being a superior predictor).

c. A very highly significant correlation (-0.78 - mean, not including GPCP-TREND) exists between STANDARD ERROR and NNS. In this study, where all the data sets are of a similar size, NNS is the most important contributing factor to the accuracy of interpolation (note, with the exclusion of a couple of outliers, the stronger common negative trends of the NNS/GPCP-STANDARD ERROR scattergram - Figure 6.28a). Interestingly, where the interpolation methods

produce generally poor results (SYMAP, GHOT and GPCP-TREND), this correlation coefficient is markedly poorer, underlining the %BAD/NNS coefficient in suggesting that extreme error is not necessarily a direct result of NNS.

In addition, the reactions of the individual data set types to particular interpolation methods may be examined by considering the mean consistency, precision, accuracy and bias statistics for each data set type. Individual data set means were generated from each of the relevant data sets for each of the interpolation methods, for consistency (ORIGINAL MIN/MAX), precision (KURTOSIS), accuracy (STANDARD ERROR) and bias (MEAN). The results are summarised in Table 6.25.

Several general conclusions may be drawn from Table 6.25 in relation to the general success of particular data sets within the interpolation methods:

a. Interpolation accuracy and consistency are directly correlated with NNS. As the NNS decreases (GRID, CONTOUR, SCATTER, BREAKLINE, RIVER), so accuracy decreases.

b. Interpolation bias is strongly associated with data set type. The surface-specific data sets generate predominately negative systematic error - RIVER (strongly negative), CONTOUR (negative) and BREAKLINE (weakly negative). Alternatively, the surface-random data sets generate predominately low or positive systematic error - SCATTER (positive) and GRID (low error).

METHOD	BREAKLINE				CONTOUR			
	accuracy	bias	consistency	precision	accuracy	bias	consistency	precision
S11-N	3.13	0.57	17.48	3.48	2.48	0.34	12.77	3.15
S11-Q	3.03	0.60	15.85	3.12	2.49	0.45	12.19	2.97
S11-O	3.22	0.65	15.99	2.96	2.71	0.57	13.05	2.95
S11-V	3.07	0.64	17.44	3.25	2.52	0.39	12.26	3.01
SYMAP	3.79	0.10	22.36	3.16	3.28	0.46	16.84	2.93
GPCP	3.23	0.19	30.26	4.33	2.13	0.19	15.09	3.20
GINOSURF	2.49	0.02	63.70	4.24	1.94	0.08	25.48	3.07
GHOT	4.32	0.41	159.50	8.09	4.81	0.43	34.76	3.48
KRIGE	2.11	0.11	11.03	3.41	1.70	0.13	10.20	2.84
MINCURV	29.78	5.95	151.17	5.83	16.95	4.62	136.18	8.32
SACM	2.32	0.05	15.06	4.26	1.93	0.00	11.67	3.11
BRAILE	3.35	0.31	17.21	3.11	3.15	0.08	16.05	2.93
MULTI	2.05	0.12	10.09	3.47	1.69	0.10	9.50	2.85
GPCP-TREND	6.02	0.02	22.73	3.16	6.00	1.22	17.98	2.66
MEAN NNS	1.122				1.350			

Table 6.25a Summary of mean bias, consistency, accuracy, precision for each data set type for each interpolation method.

(Superscripts denote the number of extreme values omitted from the meaning process due to resultant extreme statistical distorti

METHOD	GRID				RIVER				SCATTER			
	accuracy	bias	consistency	precision	accuracy	bias	consistency	precision	accuracy	bias	consistency	precision
SII-N	1.70	0.08	11.57	3.38	5.93	3.18	33.40	4.16	2.68	0.47	15.12	3.46
SII-Q	1.71	0.10	11.19	3.28	6.09	3.41	28.15	4.61	2.52	0.42	15.94	3.38
SII-O	1.85	0.16	11.51	3.38	5.75	3.56	26.53	4.19	2.59	0.49	15.40	3.17
SII-V	1.71	0.10	11.39	3.27	6.85	3.16	35.53	4.96	2.52	0.39	16.41	3.41
SYMAP	3.12	0.37	16.93	3.16	6.13	2.64	39.99	4.27	3.34	0.12	20.38	3.22
GPCP	1.34	0.13	10.25	3.66	8.46	2.56	51.82	3.06	2.36	0.16	20.84	3.90
GINOSURF	1.31	0.00	10.24	3.88	6.66	2.86	108.56	5.04	2.11	0.29	20.06	3.64
GHOT	1.33	0.00	10.11	4.06	10.60	4.87	204.40	6.55	2.32	0.26	49.92	5.86
KRIGE	1.44	0.01	11.08	4.74	5.44	3.03	19.40	3.70	1.77	0.19	12.13	3.38
MINCURV	2.76	0.05	46.84	9.31	67.41	3.05	256.96	4.82	8.50	0.64	73.66	7.90
SACM	1.64	0.04	11.71	5.47	5.36	2.10	20.54	3.79	2.16	0.29	13.82	4.23
BRAILE	2.78	0.05	16.25	3.03	4.91	4.11	66.18	3.62	3.14	0.25	16.36	2.88
MULTI	1.31	0.02	9.28	3.81	5.57	3.82	18.94	3.37	1.74	0.31	11.87	3.35
GPCP-TREND	7.26	0.16	26.70	2.39	20.82	3.59	43.95	2.84	6.95	0.07	25.60	2.62
MEAN NNS	2.052				0.786				1.180			

Table 6.25b Summary of mean bias, consistency, accuracy, precision for each data set type for each interpolation method. (Superscripts denote the number of extreme values omitted from the meaning process due to resultant extreme statistical distortion)

c. Precision displays a similar pattern to accuracy and consistency, although GRID and CONTOUR are interchanged. In general, precision decreases as accuracy decreases and the NNS statistics decreases. This is most likely to be a result of a greater error range in the lower NNS data sets.

In addition, more specific conclusions may be drawn for the various interpolation methods:

a. The data-averaging methods (SURFACE II GRAPHICS) have good consistency of data set type response, although BREAKLINE and RIVER data sets produce a greater variation of consistency and accuracy. This is a direct result of the NNS, since these data sets have large data voids which are much more susceptible to changes in the search method utilised when simple data-averaging functions are being utilised. Only the RIVER data sets produce a negative systematic error component.

b. The data surface-averaging processes (SYMAP and GPCP) both produce a marked increase in negative systematic error for all data sets. While in SYMAP, all data sets have a consistent deterioration in accuracy, consistency and precision, GPCP shows a marked deterioration in accuracy and consistency only for the sub-1.150 NNS data sets (the strongly lineated data sets), and a marked increase for the

super-1.150 NNS data sets.

c. The performance of the pointwise quadratic methods (GINOSURF and GHOT) is related primarily to the surface-specific/random nature of the data sets. Use of these techniques with the surface-random data sets (GRID and SCATTER) produce moderate increases in accuracy and consistency over the previous methods. The lineated surface-specific data sets (BREAKLINE, CONTOUR and RIVER) produce strong decreases in accuracy and consistency within GHOT. This is associated with the fact that the lineated data sets produce ill-conditioned normal equations in the node interpolation, exaggerated where the data distribution is poor, i.e. in low NNS data sets, and where the search method finds insufficient points.

While GINOSURF produces low magnitude bias, and GHOT produces a higher magnitude bias, all bias follows the general trend of all the interpolation methods - use of SCATTER data leads to a positive bias, use of GRID data gives no systematic error component and the surface-specific data sets all produce negative bias.

d. The accuracy of interpolation by KRIGE is directly correlated to NNS. Use of all data sets gives similar consistency, with the exception of RIVER data, and this

would suggest that only data sets with an NNS greater than 1.0 are suitable for this technique.

e. MINCURV, while producing generally unpredictable results, is much more suited to point-based, surface-random data sets without appreciable data voids.

f. SACM and MULTI generate superior interpolated results for all data sets, although, as might be expected, data sets with generally lower accuracy results by other methods have the largest increases in accuracy. These two methods (SACM and MULTI) are of most benefit with data sets which the other methods cannot cope with satisfactorily.

g. BRAILE and GPCP-TREND, which produce generally low accuracy results (i.e. they are generally poor interpolators), show most consistency between varying data sets. Nevertheless, use of RIVER data produces very poor results and therefore data sets with an NNS less than 1.0 should not use these methods. Interestingly, results from the GPCP-TREND interpolated GRID results are of a lower accuracy than are those from the BREAKLINE, CONTOUR and SCATTER data sets, and this emphasises that these two procedures do not obey any general data sets type convention, but produce a large amount of surface-associated error.

CHAPTER 7

GRID-TO-ISARITHM INTERPOLATION

7 ISARITHM INTERPOLATION

7.1 INTRODUCTION

Having considered the effects of the first stage in isarithm interpolation, this Chapter will evaluate the effects of the second stage - grid-to-isarithm interpolation (also termed 'contour threading'). As in the previous chapter, the accuracy of individual methods of interpolation will be considered and the relationships between interpolation error and data characteristics studied. A similar investigation will be made of triangulation-to-isarithm interpolation, thus allowing a comparison of the two basic methods of generating isarithms on the computer - by triangulation and by grid interpolation.

7.1.1 Résumé Of Previous Research

Unlike random-to-grid interpolation, a very limited amount of research has been performed on evaluating grid-to-isarithm software. Excluding descriptive literature on the implementation of various software (e.g. Batcha and Reese, 1965; Heap and Pink, 1969; Walters, 1969 and Rhind, 1971), studies have been limited to individual programs and their products. These range from the folksy, where resultant isarithms are compared graphically with some 'benchmark' (e.g. Royal College of Art, 1972 - considering the effects of using the ECU and SACM packages), through restricted package

evaluation (e.g. Schilcher, 1977 and Assmus and Stanger, 1978 - considering SCOP by using graphical overlays and deriving Koppe accuracy statistics), to system and package comparison (e.g. Blachut and van Wijk, 1976 and Gustafson, 1977 comparing different systems of orthophoto contour generation using adjacent plots and Koppe accuracy statistics). If Koppe statistics are examined, then frequently these are evaluated from Lindig-based theories (see 5.2.2) and compared with international map specifications - frequently for German and Swiss national series (e.g. Lips, 1964; Blachut and van Wijk, 1976; Gustafson, 1977; Schilcher, 1977 and Assmus and Stanger, 1978). Discussion in such a case revolves around whether the process (data collection to map) has attained the required standard and does not involve evaluating the general performance of the grid-to-isarithm interpolation.

There is therefore a distinct lack of firm information in the literature concerning this important stage of isarithm interpolation. Walden (1972) propounded one of the few theories. While avoiding any practical examination of the techniques, she suggested that:

> "Obviously, the second step of the automatic contouring algorithm is entirely dependent on the first step. The interpolation procedure used to connect all points of equal value (or evaluation) on the surface uses the grid values, and not the raw data, for reference. Furthermore ... contour line smoothness is a function of the fineness of the grid. Thus, both the accuracy and the appearance of the contour map depend on the precision of the grid from which it is drawn" (Walden, 1972, 5).

Similarly, there is a distinct lack of information concerning the relative abilities of interpolating isarithms

utilising a triangulation net and utilising a regular grid. Frye (1978) proposed methodology to be used to compare the performances of the two methods - Triangulated Irregular Network (TIN) and Uniform Rectangular Grid (URG). This involves applying the theory of Court (see 5.3.2) to compare the results of precise applications using both methods, with the 'known truth'. Walden (1972) proposed a general 'rule-of-thumb' for using the triangulation or grid methods based on Switzer, Mohr and Heitman (1964). She suggested that, if autocorrelation amongst grid values is high, then a triangulation net should be used, since the average distance from a point being estimated to a grid node will be smaller. However,

> "When the correlation is low, the square grid in theory provides a slightly better estimate of the surface. This is because the introduction of a fourth point should eliminate some of the noise in the data used to make the estimation" (Walden, 1972, 6).

Clearly there is scope for detailed analyses of both the performances of the grid-to-isarithm algorithms and also the relative accuracies attainable from data sets utilising a triangulation-to-isarithm algorithm.

7.1.2 Brief Résumé Of Methods Utilised

Seven isarithm interpolation routines were considered in this aspect of the research. The self-written program 'DTC' (see A1.4.6) was the only triangulation-to-isarithm program used and performed a linear interpolation of isarithms without smoothing - this aspect was considered within grid-to-isarithm

interpolation. Six grid-to-isarithm packages/subroutine libraries were utilised with the following summarised options:

a. SACM - the isarithms are threaded by linear interpolation through each grid cell;

b. SURFACE II GRAPHICS - the isarithms are threaded through each cell by linear interpolation and smoothed;

c. GHOST - using the diagonal, each cell is divided into two triangles and the isarithms are linearly interpolated without subsequent smoothing;

d. CONSYS - a centre point (mean of four surrounding nodes) is derived for each cell, thus establishing four triangles through which the isarithms are threaded by linear interpolation without any subsequent smoothing;

e. GINOSURF - a centre-point (mean of four surrounding nodes) is derived for each cell, thus establishing four triangles through which the isarithms are threaded by linear interpolation and subsequently smoothed;

f. GPCP - the resolution of each cell is increased from a patchwise third order surface fitted to the immediate

neighbouring nodes - isarithms are subsequently threaded by linear interpolation through the resultant sub-cells.

7.1.3 Evaluation Procedure

Within this research, isarithms were interpolated from three different resolutions of observed regular grid for each of the three models, using five different packages. Only the finest resolution grid for each model was considered in SACM. In addition, isarithms were directly interpolated for each of the non-grid point data sets using a triangulation net. These two aspects will be discussed separately.

As a result of previous discussion (Chapter 5), three methods were used for evaluating the accuracy of the isarithms derived from the various methods and grid resolutions.

For each grid resolution and model (point data), overlay maps were prepared incorporating the products of all the interpolation methods with the reference isarithms to give a graphical representation of both geometric accuracy and morphological trueness. In order to avoid any ambiguities, a maximum of three interpolated isarithm maps were overlain with the relevant reference map using different coloured inks. The reference isarithms were assigned red and the other methods (grid) and data sets (point data) assigned different colours and line types (see Table 7.1a and 7.1b). Unfortunately, the limited number of different inks available meant that for each data set two adjacent overlay maps (see Figures 7.1 to 7,9 and

Table 7.1a - Grid colour/line conventions

METHOD	COLOUR	LINE
CONSYS	green	pecked
GHOST	blue	pecked
GINOSURF	blue	solid
GPCP	black	solid
SACM	black	pecked
SURFACE II	green	solid

Table 7.1b - Point data colour/line conventions

DATASET	COLOUR	LINE
BREAKLINES	black	solid
CONTOURS	blue	solid
RIVER	blue	pecked
SCATTER	black	pecked

7.19 to 7.21) were required to display all the methods. On each map, the box beside the title is equivalent in size to a grid-cell.

Geometric accuracy was established using the theory of Court (see 5.3.2) within the program 'CONCOR' (see A1.4.4). Summary diagrams were prepared from the output of CONCOR and separated according to model (see Figures 7.10 to 7.12 and 7.22 to 7.24). Within each diagram, the isarithm correlation for each level is displayed in the top graph, with an average for the map in the top corner. The middle graph summarises the total percentage area of the map within each isarithm level and the lower graph summarises, for each level, the percentage area inside the reference isarithms but outside the interpolated isarithms (positive), and the area inside the interpolated isarithms but outside the reference isarithms (negative). For each model, the full grid (finest resolution) is shown by a solid line, the half-grid (medium resolution) by a pecked line and the quarter grid (coarse resolution) by a finely pecked line.

The theory of Lindig (see 5.2.2) was also used to establish both geometric accuracy and morphological trueness. Summary diagrams were prepared for each model showing the position, elevation, direction and curvature errors (see Figures 7.13 to 7.15 and 7.25). Unique symbols are associated with each colour and line type which were assigned similarly to the overlay map comparisons. Within gridded data comparisons, the horizontal axes increase with grid size and increasing resolution. Within point data (triangulation) comparisons, the leftmost points

represent INCH data, the middle points represent LIAN data and the rightmost represent FORV data.

In addition, Figures 7.16 to 7.18 and 7.26 summarise the Koppe accuracy statistics with (in red) the Swiss 1:10,000 (solid) and German 1:5,000 (pecked) map accuracy specifications.

Overall, the methods applied represent a detailed and exhaustive appraisal of both geometric and morphological accuracies.

7.2 GRID-TO-ISARITHM

7.2.1 Graphical Evaluation

In the random-to-grid comparison, a numerical product was generated, necessitating primarily a numerical or statistical evaluation. In the grid-to-isarithm comparison, a graphical product is to be considered thus requiring a preliminary graphical evaluation as a pre-requisite to a later statistical analysis. The isarithm maps will be considered primarily in terms of grid resolution, and within this, the effects of surface roughness and individual packages will be evaluated. The success of the packages will be considered in terms of their ability to regenerate macro- and micro-features, the general smoothness of the resultant isarithms, any resultant ambiguities, their discrepancies with other methods and their closeness to reality. Throughout the example, numbers refer to

the numbered circles on the respective figures.

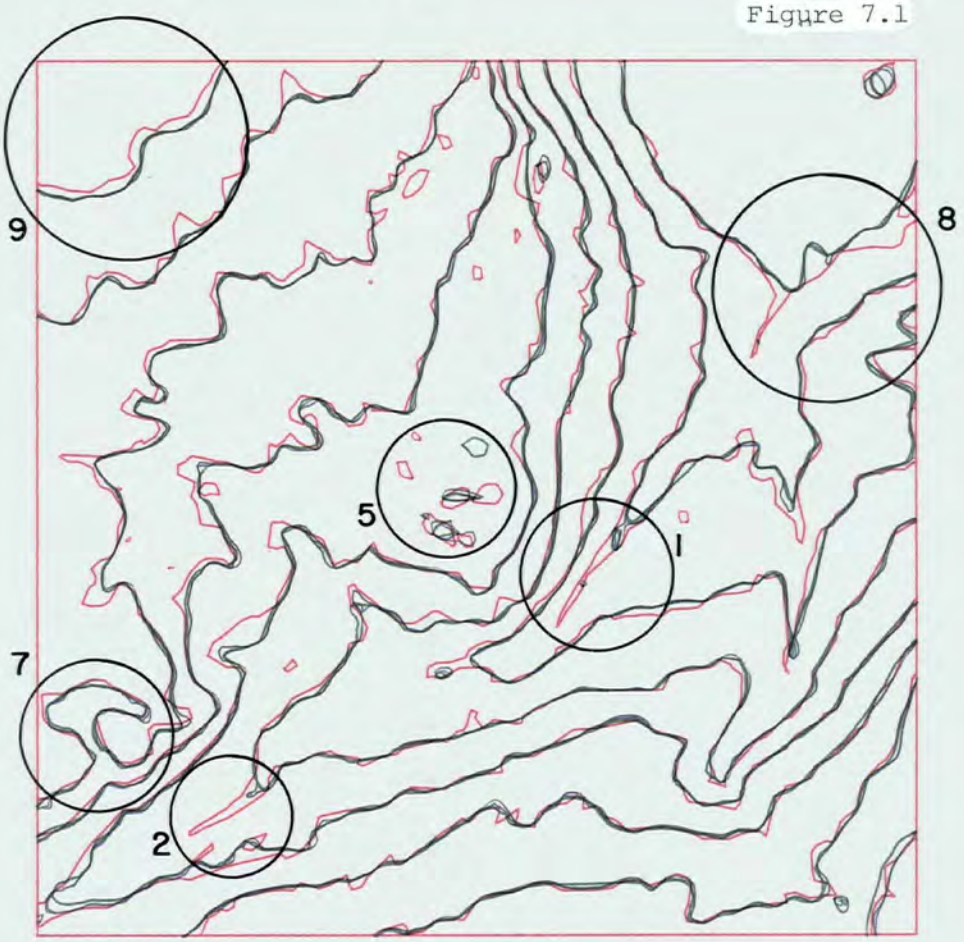
7.2.1.1 Full-grid - Finest Resolution

The full-grid maps are shown in Figures 7.1 (INCH), 7.2 (LIAN) and 7.3 (FORV).

In general, all maps show a good representation of all macro-features. Use of the smooth, highly autocorrelated LIAN surface (see Table 4.3) produces the best isarithm interpolation of such features. INCH has problems in the main valley representation (1 and 2) which are a direct result of the feature width being less than the grid resolution. FORV has problems (see 3), where the interpolated isarithms are distinctly different from the reference isarithms. This results from the grid nodes being slightly lower than the local reference isarithms as a result of the accuracy of the sampling process (see 3.4.4). The error is accentuated due to the flatness of the terrain at this point.

At the macro-level use of FORV data leads to good delineation of small knolls and pits (insular isarithms). However, while the surface impression is generally conveyed well, the actual product often lacks accuracy (4). LIAN, and to a greater extent INCH, are less successful due to the micro-features having a finer resolution than FORV. It is significant that those packages which operate smoothing, whether directly (GINOSURF) or indirectly (GPCP), give consistent and superior products: small looped isarithms tend to be more accurately placed and of more realistic size (5 and 6). Conversely, a more ragged and variable product is

Figure 7.1



INCH : 51X51 □

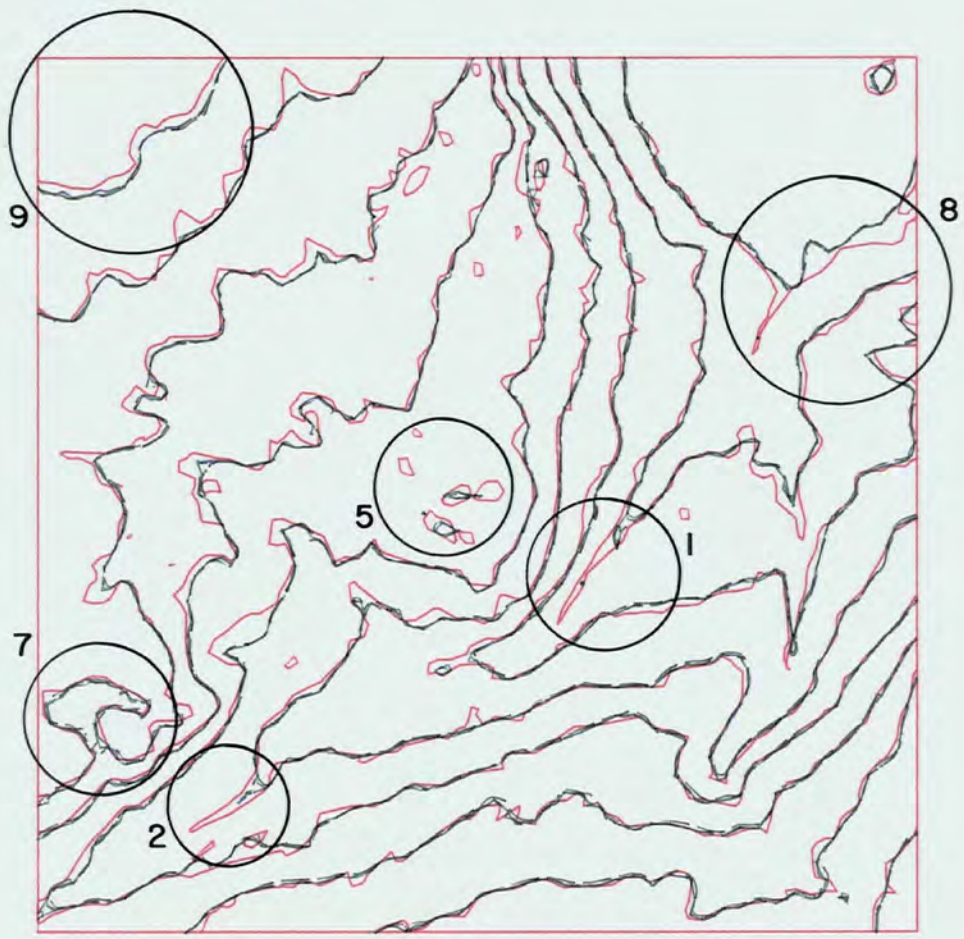
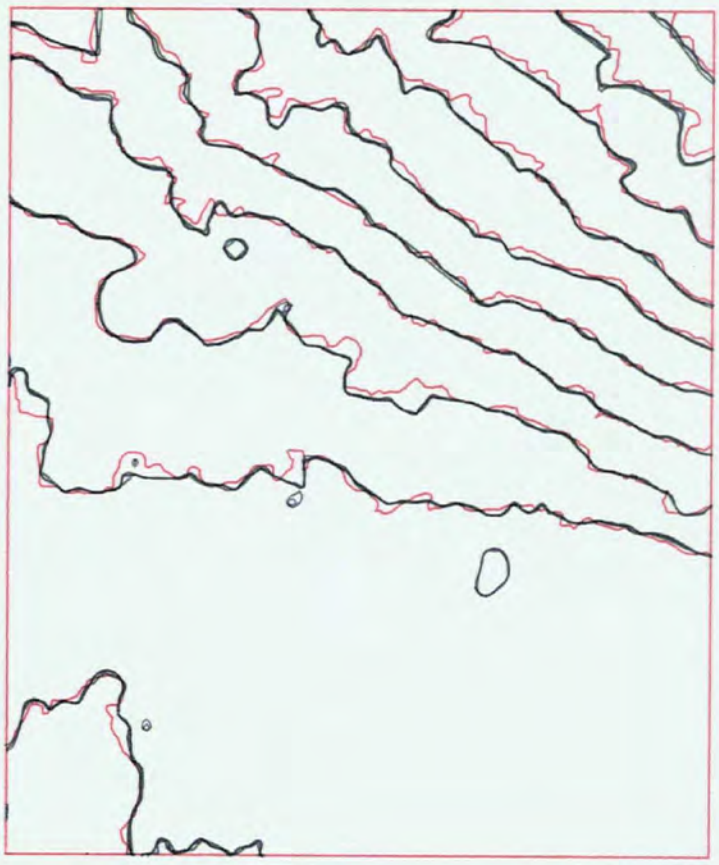


Figure 7.2



LIAN : 51X61

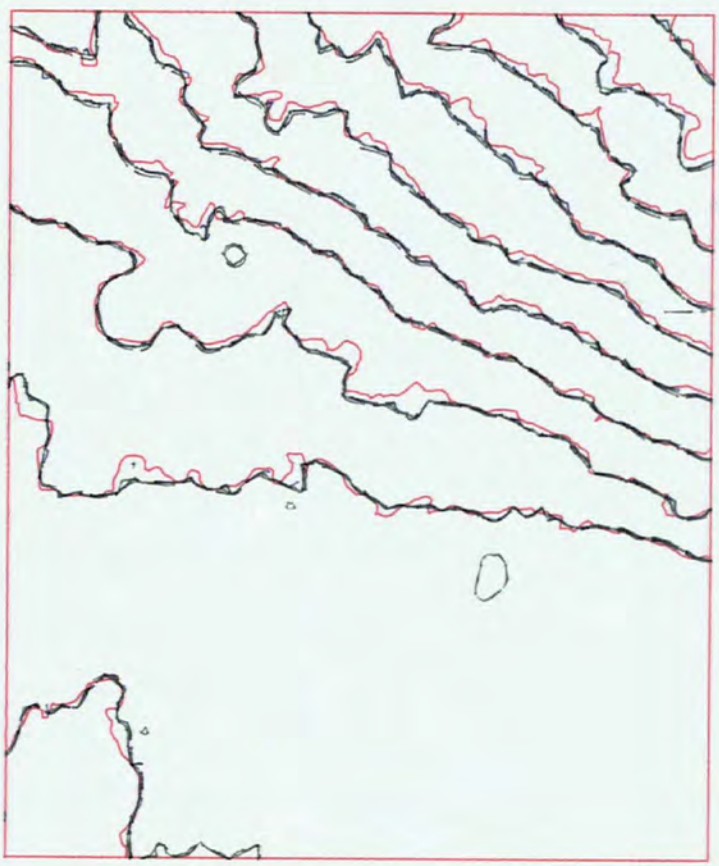
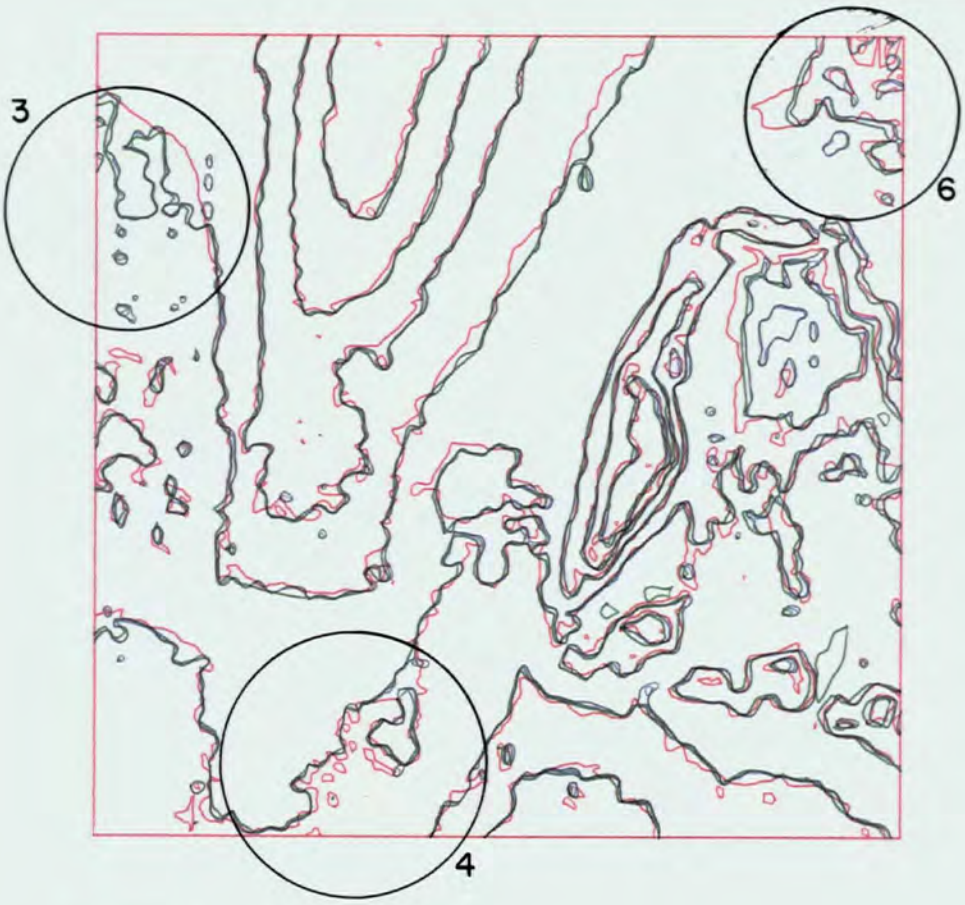
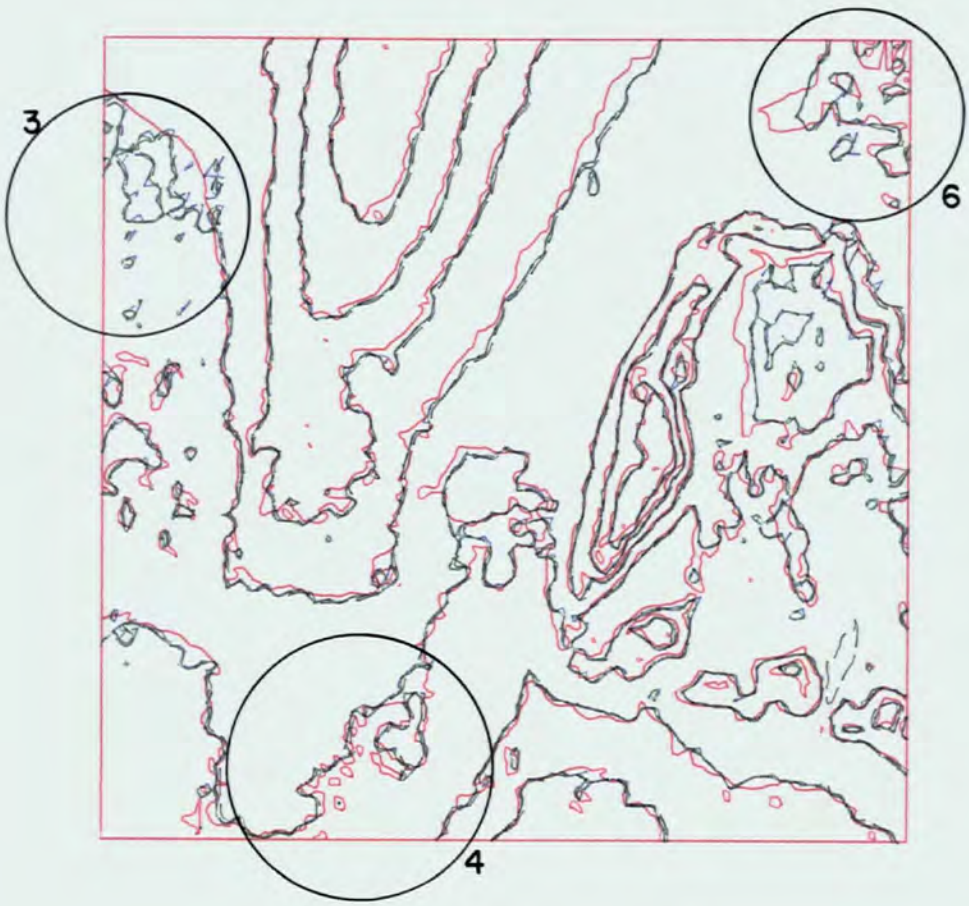


Figure 7.3



FORV : 65X65 □



generated by CONSYS and, more significantly, GHOST (5 and 6).

GHOST shows several errors (eg. the corrie at 7), although this is balanced by a superior delineation of the major river valley in INCH. These conflicting results are due to the grid-to-triangle process. Because GHOST uses the 'diagonal method' (see Figure 2.13a), the grid is not refined any further. The isarithm, for example, will not pass between two low points in the river valley which are both below the isarithm level and are at opposite corners of the grid-cell. However, the 'centre-point' methods (GINOSURF and CONSYS) generate an additional point in the centre of the cell from a mean of the four corners. If the river banks are steep, then this mean may be high, resulting in a saddle point which breaks the resultant isarithm into a continuous valley and a separate small pit (1 and 2). GPCP succeeds in modelling the river valley due to the nature of the polynomial patchwise process involved in deriving the finer resolution sub-grid. This defines the local shape of the surface better than an arithmetic mean.

As regards smoothness of the interpolated isarithms, all data sets and methods produce similarly smoothed products with the exclusion of GHOST (as previously discussed). There would appear to be no requirement for any form of polynomial smoothing after linear interpolation at this grid size and map scale.

Similarly, in these data and with the exclusion of CONSYS there are no ambiguities (for example meeting or crossing isarithms). CONSYS problems occur primarily at the small

insular loop-isarithm/line-isarithm interface (3 and 6). Unlike other packages, the isarithms are not traced through the grid and therefore grid lines are not 'checked-off' and thus prevented from being used several times. The same grid node (if it has an equivalent 'z' value to the isarithm level) may be used several times (up to four) and thus isarithms cross and ambiguities occur.

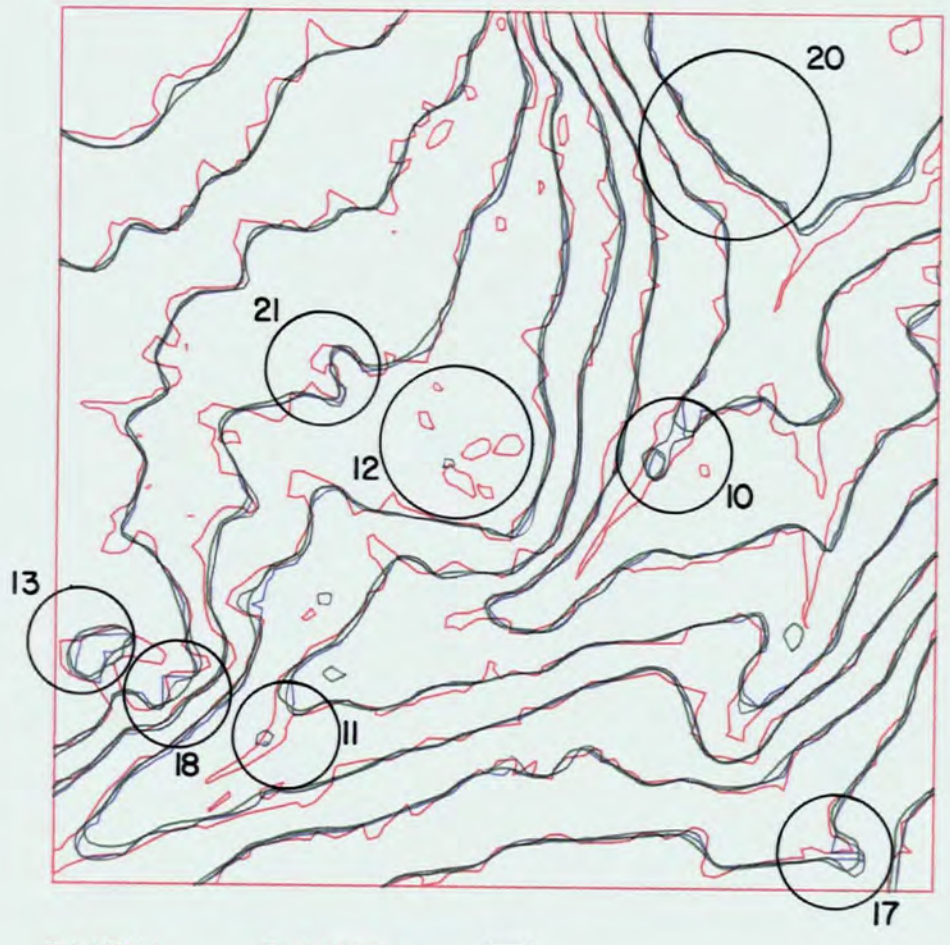
Overall, at this fine resolution, discrepancies between methods are few - excluding those previously discussed concerning GHOST and CONSYS - and are within generally acceptable plotting tolerances. The interpolated isarithms appear to replicate the surface sufficiently well, although this will be examined statistically in subsequent sections. Minor differences do occur notably in FORV (3 and 6) and INCH (8 and 9), although this is mainly in flat areas and would be expected on any isarithm map where low gradients occurred.

7.2.1.2 Half-grid - Medium Resolution

The half-grid maps are shown in Figures 7.4 (INCH), 7.5 (LIAN) and 7.6 (FORV).

As regards the delineation of macro- and micro-features, LIAN demonstrates no significant change from the full grid, primarily as a result of the high surface autocorrelation. However, there is a significant decrease in small insular isarithms, with only the major ones remaining in a reduced size. This is to be expected. As the grid size is halved, so critical nodes are lost and the wavelength of feature which may be defined is doubled.

Figure 7.4



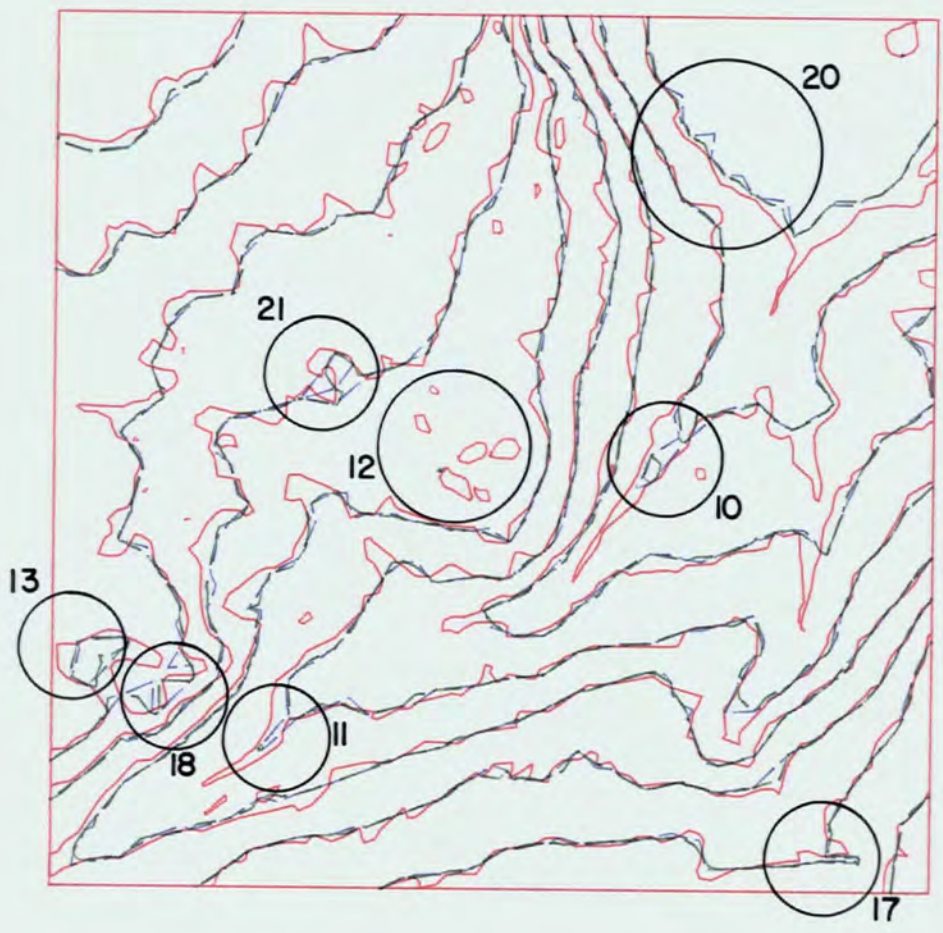
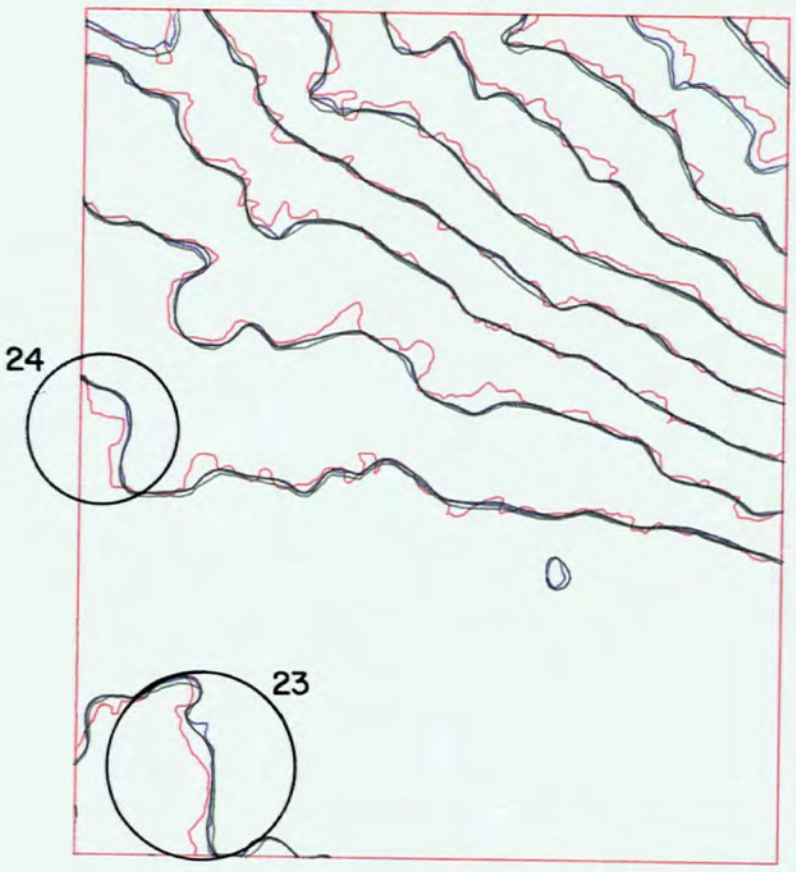


Figure 7.5



LIAN : 26X31 □

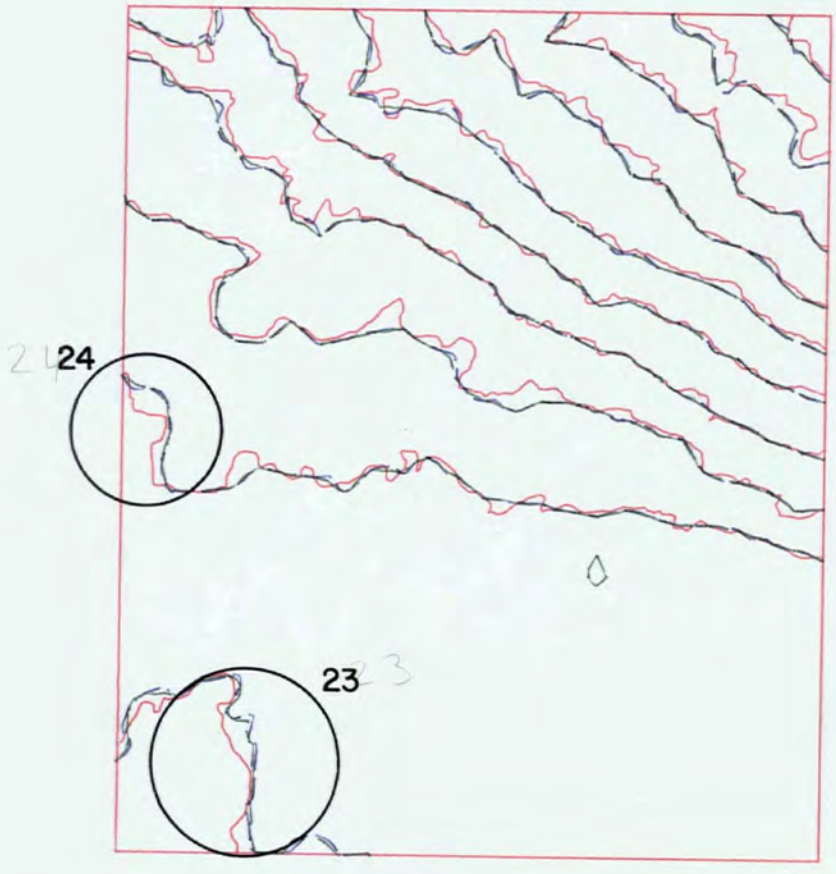
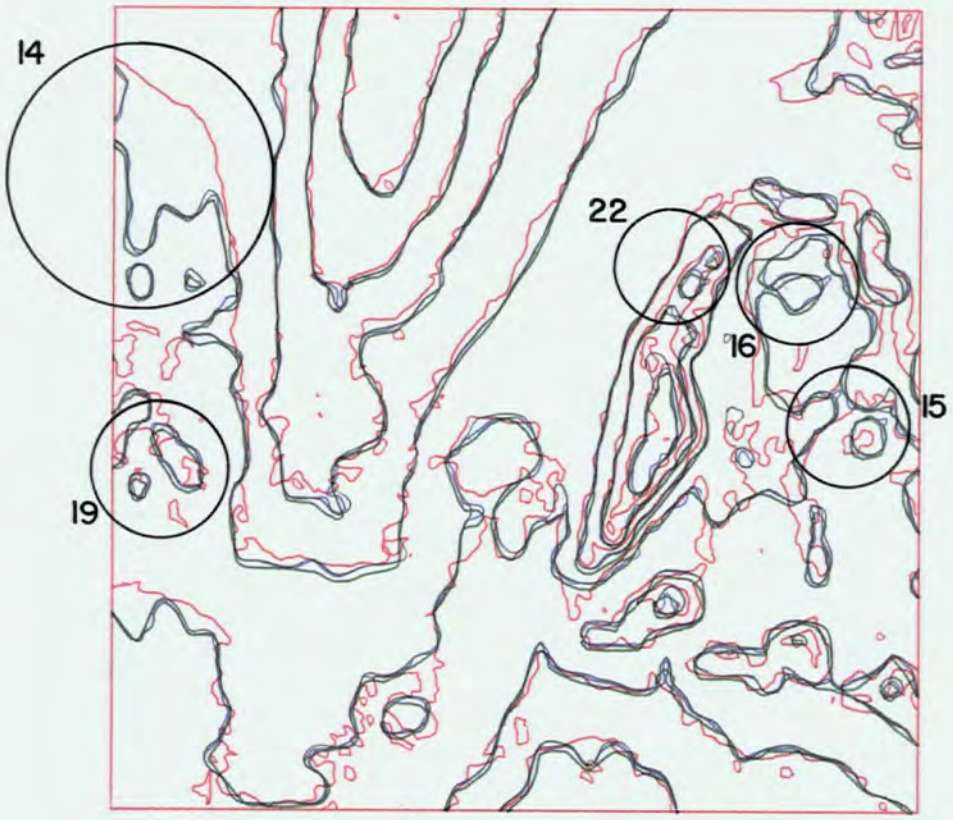
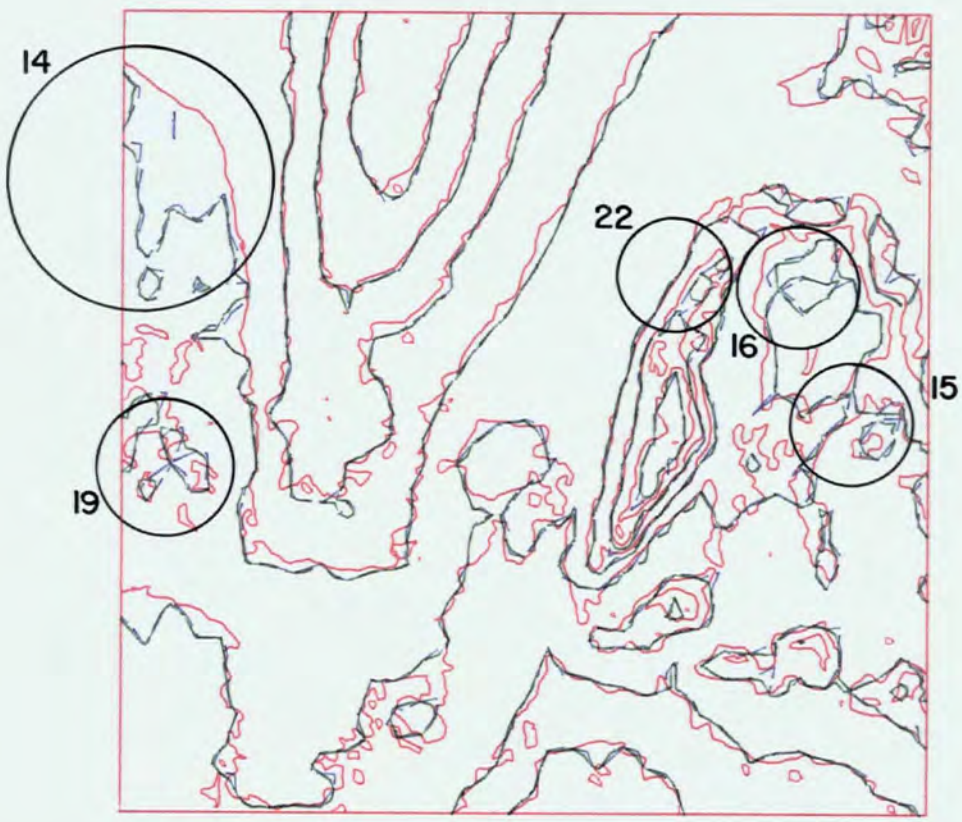


Figure 7.6



FORV : 33X33 □



The less autocorrelated INCH surface suffers from further river valley recession as a result of the coarser grid resolution. An associated effect is that the grid-to-isarithm method becomes more critical. Opposite grid corners are further apart and therefore the 'saddle point' problem, derived from the 'centre point' grid-to-triangulation techniques, causes more distinctive variations between GPCP and GHOST and the other methods (10 and 11). As with LIAN, there are virtually no insular isarithms with the exclusion of those (e.g. 12) which are caused by vital grid nodes being lost in decreasing the grid resolution.

FORV best illustrates the major fluctuations of shape that may occur with halving the grid resolution. In 14 and 15 considerable changes occur in flat areas where vital nodes are lost and the effect of the remaining ones is increased as the distance between nodes has increased. Significantly, GPCP produces slightly different results from the other techniques, as the patchwise polynomial function used may increase further the effect of the remaining nodes (16).

It is at this resolution of grid and map scale that the influence of a smoothing algorithm affects the results. Once again, the highly autocorrelated LIAN shows minimal differences. FORV (16) and INCH (17 and 18) not only show considerable variance between the smoothed isarithms and the angular isarithms, but the effect of the coarser triangulation within GHOST is apparent (FORV-19 and INCH-18 and 20). Excluding LIAN, ambiguities are more noticeable as a result of smoothing (16), with coarser triangle resolutions (INCH-21 and

FORV-22 - GHOST) and the previously mentioned problem with CONSYS due to the scan production of isarithms (16). With regard to the consistency of methods and their ability to replicate reality, there is a low variance among all LIAN interpolated isarithms. Where isarithms are smooth, a slight 'downhill' slide takes place (23) which is balanced by a slight 'uphill' movement where isarithms are more angular (24). Within INCH a similar consistency is noted away from the corrie and river areas previously discussed. FORV shows the greatest variance as a result of the smaller range of altitude and gradient which has its maximum effect on isarithm position in linear interpolaton. Significantly, most variations are in flat areas.

7.2.1.3 Quarter-grid - Coarse Resolution

The quarter-grid maps are shown in Figures 7.7 (INCH), 7.8 (LIAN) and 7.9 (FORV).

At this coarsest resolution, only macro-features are maintained, with a total loss of micro-features. The effects of utilising diagonal and centre-point methods of triangulation are more marked (INCH-25 and -26), with GHOST generally showing superior interpolation.

As regards smoothness, notable differences between the various methods are seen for the first time within LIAN (27) at this resolution. While the three smoothing algorithms illustrate considerable over-generalisation, the CONSYS and GHOST methods show minor variance. Naturally, this is emphasised most in the more sensitive INCH and FORV models.

Figure 7.7

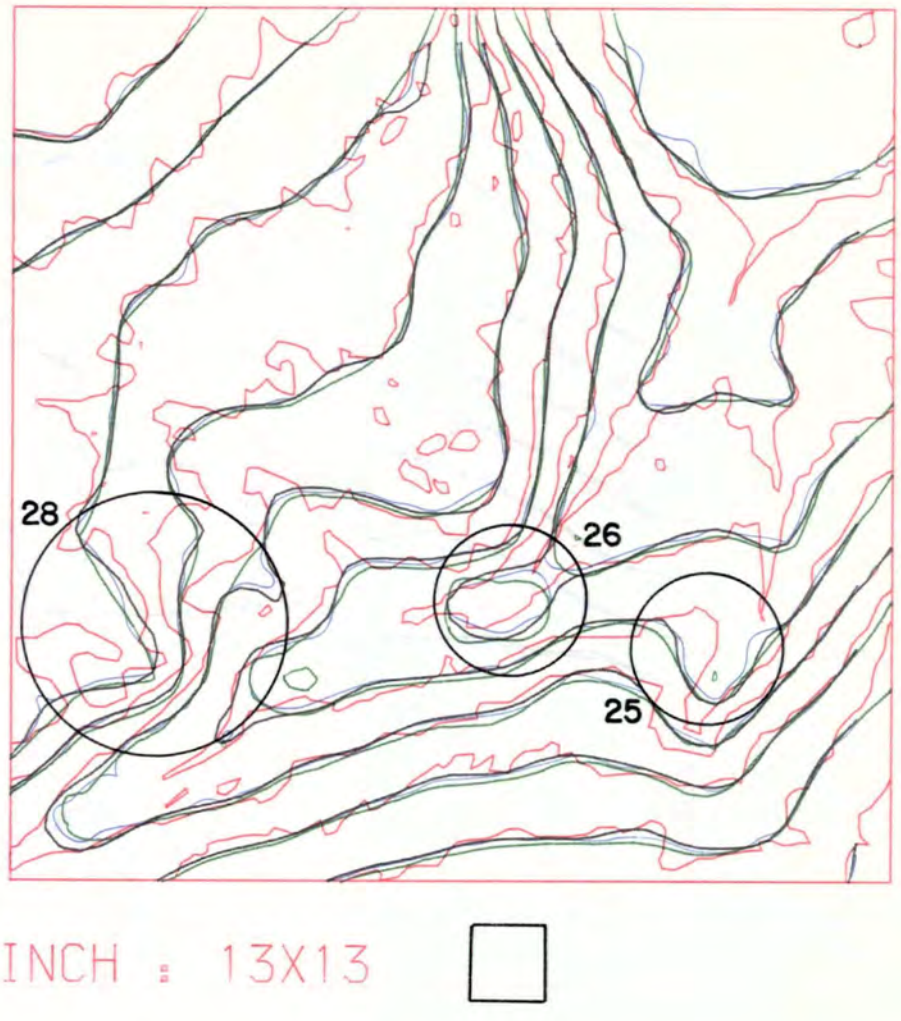
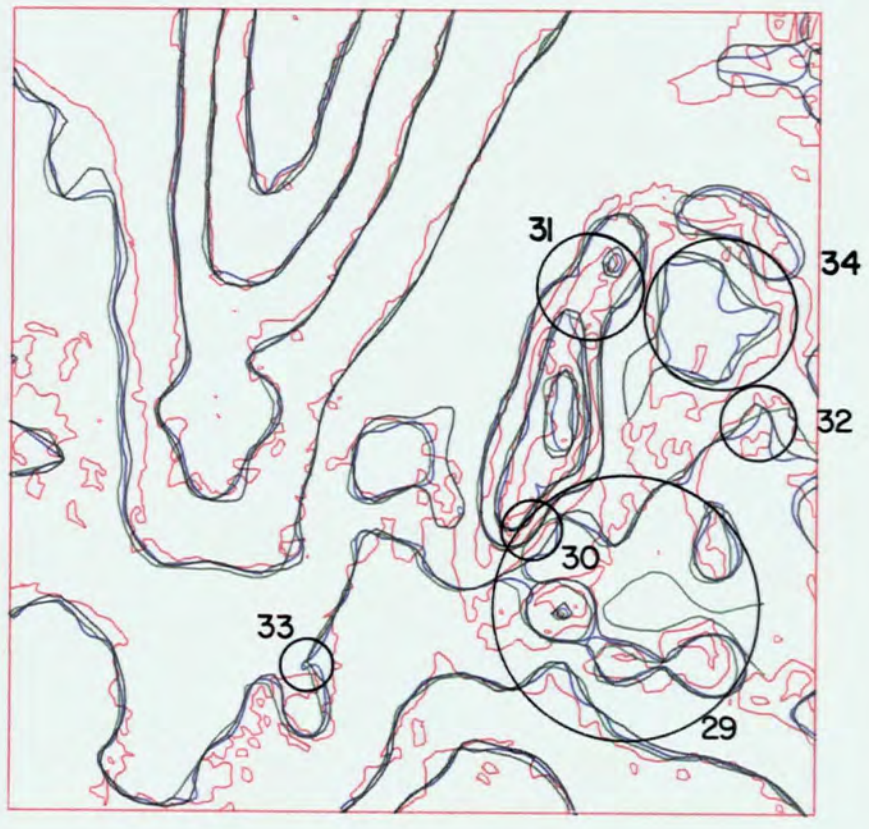
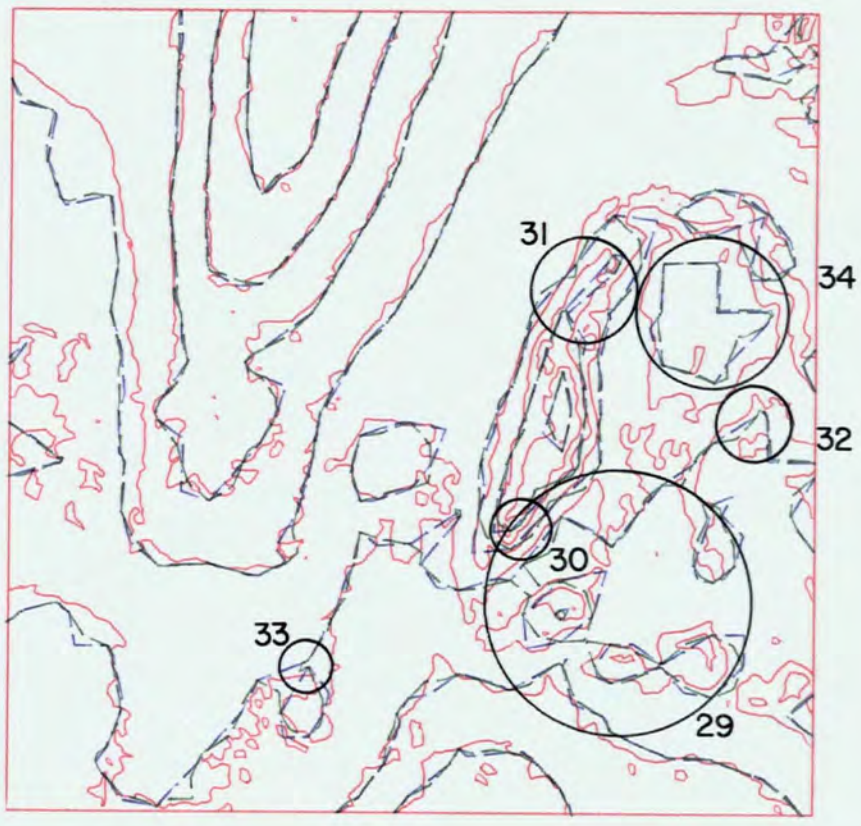


Figure 7.9



FORV : 17X17 □



The inherently simple nature of the macro-features in INCH results in greatly over-generalised isarithms with a large inter-algorithm variance (INCH-28). This is more acute in FORV, leading to severe ambiguities with CONSYS isarithms and overlapping (SACM - 30) or touching (GHOST - 31) isarithms. 'Cusping' (32 and 33) emerges only in the FORV grid. When linearly interpolating, this is caused by an isarithm changing direction sharply to encompass a grid node (see CONSYS). When this isarithm is subsequently smoothed a cusp results (see GINOSURF). As this is a feature of post-interpolation smoothing within a coarse grid it does not exist in function-based isarithm interpolation (see GPCP).

'Rectangular' isarithms (FORV-34) may exist in unsmoothed, non-function based isarithm interpolation. They occur where the grid resolution is coarse (cf. FORV-17 resolution) and where a flat grid cell(s) is surrounded by very high or low grid nodes. When linear interpolation takes place, the isarithms 'hug' the sides of the grid cells. This may additionally result in cusping if the isarithms are smoothed (see GINOSURF).

A purely graphical and descriptive analysis of the results allows one to understand some of the nuances of alternative techniques. However, it does not give precise estimates of performance; for this numerically evaluated statistics must be examined.

7.2.2 Geometric Accuracy

The geometric accuracy of the various grid-to-isarithm

methods will be discussed with reference to the correlation statistics as derived from CONCOR (see A1.4.4), position and elevation errors as derived from CONLIN (see A1.4.5) and Koppe formulae derived from the latter. These results are summarised as follows: for correlations in Figures 7.10 (INCH), 7.11 (LIAN) and 7 12 (FORV); for position and elevation errors in Figures 7.13 (INCH), 7.14 (LIAN) and 7.15 (FORV) and for Koppe statistics in Figures 7.16 (INCH), 7.17 (LIAN) and 7.18 (FORV).

7.2.2.1 Correlation Statistics

The results shown in Figures 7.10 to 7.12 confirm the previous discussion. All summary correlations for LIAN are higher than for any method with any other data set. The full-(51) and half-grids (26) perform equally well and may be considered identical - the only serious discrepancies occurring at the quarter-grid (13) size. While GPCP is consistently superior at all three grid resolutions, this is only minimally so. GINOSURF and CONSYS are of slightly lower accuracy and GHOST has the poorest accuracy. This would suggest that, amongst packages, accuracy is greater for those which employ finer intermediate resolution. Examining INCH grids, the lower surface autocorrelation means that even at the full-grid resolution the best correlations are only 0.980. As the resolution becomes coarser, the isarithm correlation decays more rapidly. While the half-grid (26) correlation coefficients may be sufficiently high and close to the full-grid (51) coefficients, the quarter-grid (13) is outside acceptable limits. Significantly, GPCP again gives

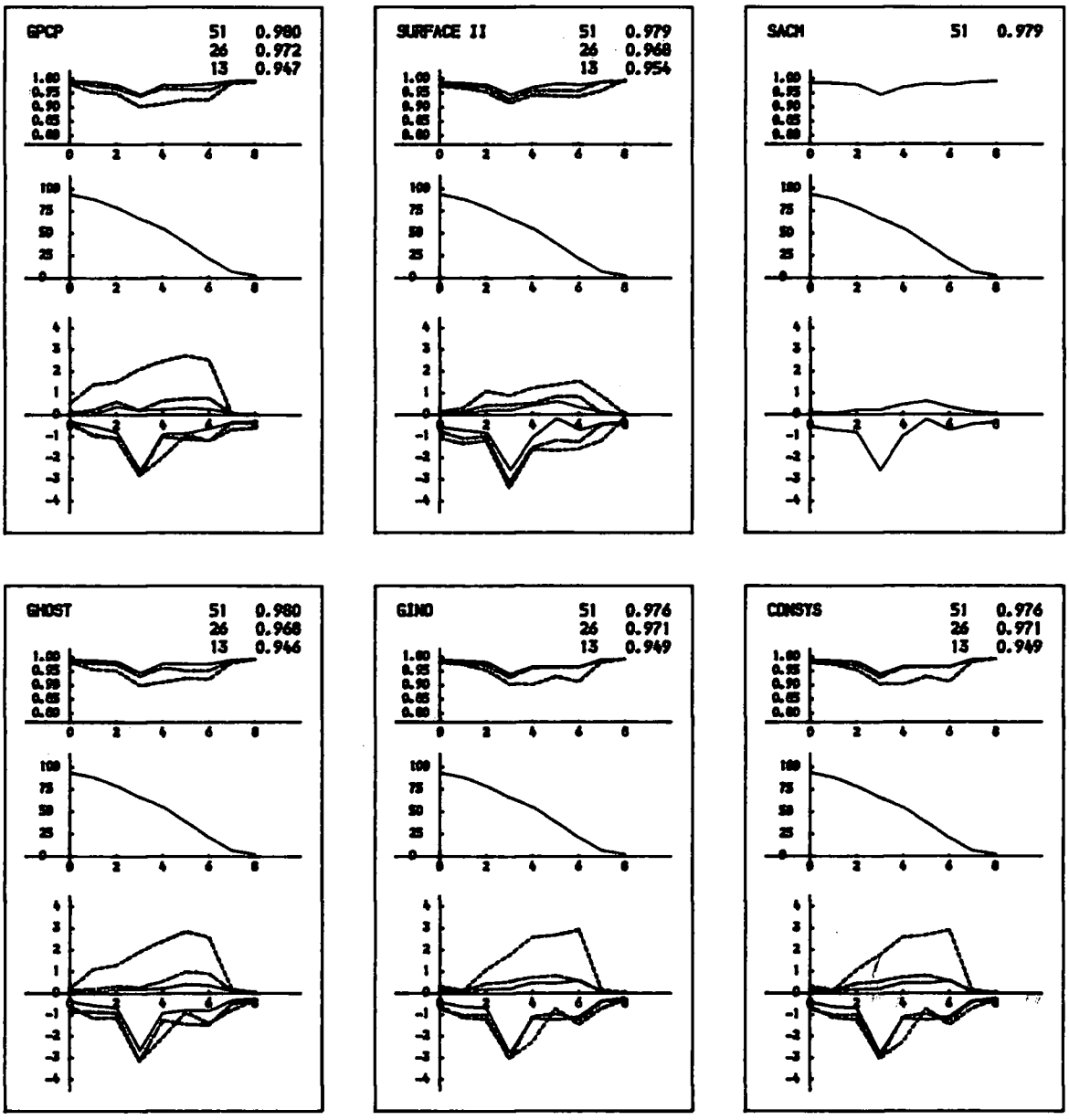


Figure 7.10 Grid-to-isarithm correlation statistics (INCH)

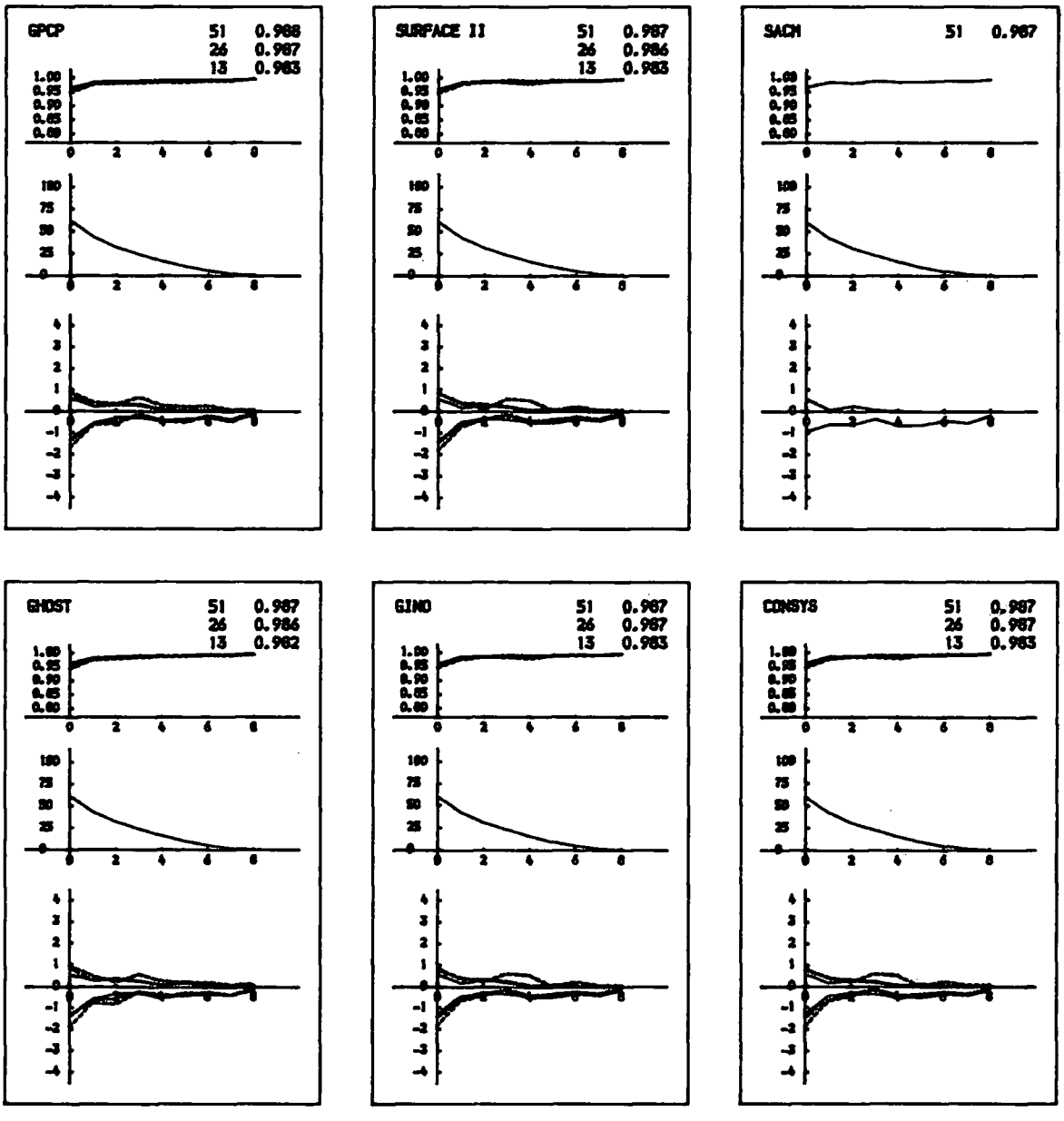


Figure 7.11 Grid-to-isarithm correlation statistics (LIAN)

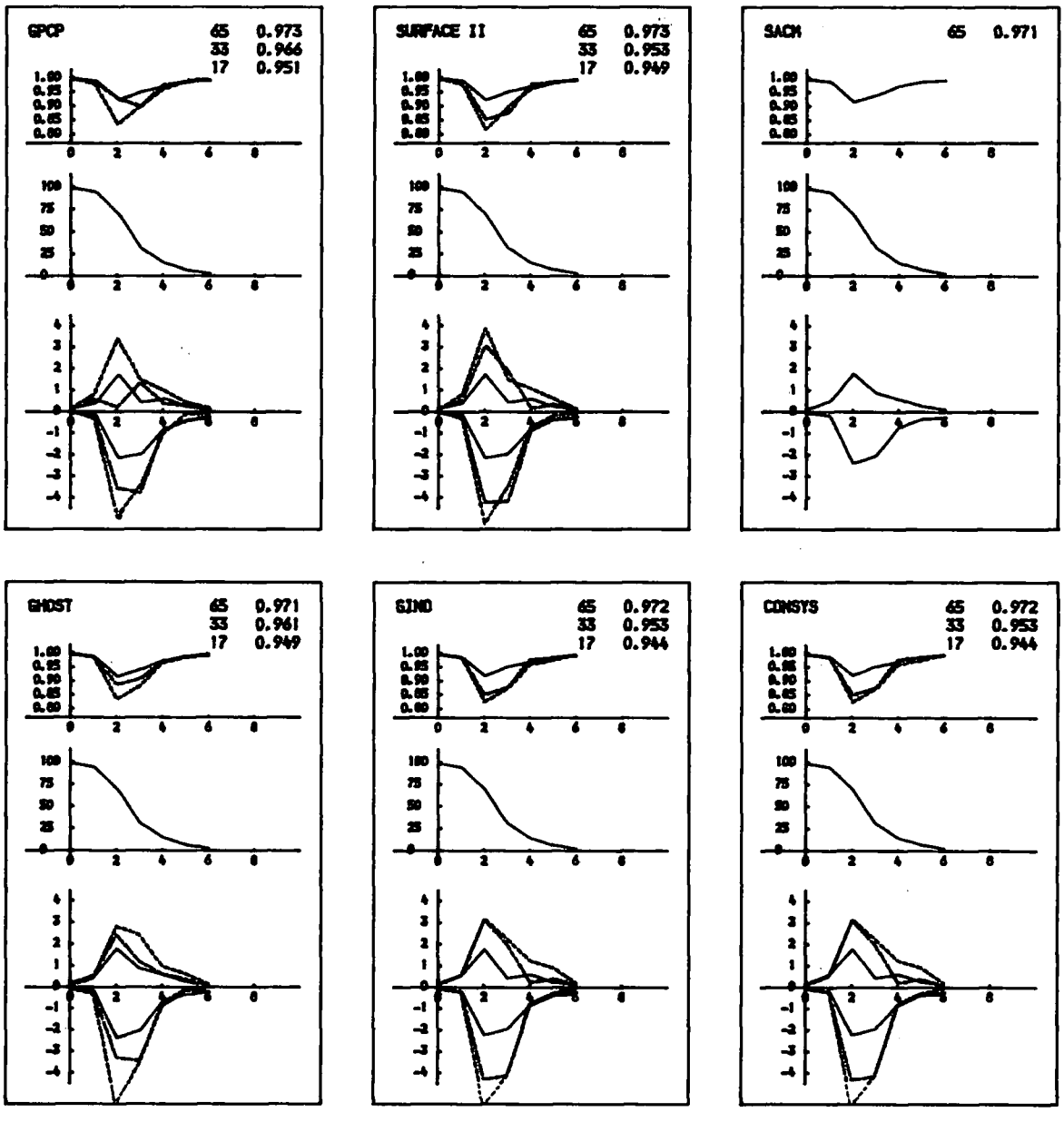


Figure 7.12 Grid-to-isarithm correlation statistics (FORV)

consistently superior results. Excluding GHOST-51 and SURFACE II GRAPHICS-13, GINOSURF and CONSYS again provide a consistent close second.

Since Table 4.3 suggested that FORV had the lowest surface autocorrelation, the poorest accuracies with the greatest variance might be expected with FORV. The finest grid resolution has the lowest full-grid isarithm correlation coefficients and there is a sharp decay to unacceptable half- and quarter-grid coefficients. This supports previous graphical evidence which suggested that only the full-grid produced a suitable product. GPCP again produces consistently the best isarithms, although GHOST and SURFACE II GRAPHICS are superior at the coarser resolutions - the benefits of a 'diagonal' triangulation have been suggested earlier in areas of low autocorrelation.

In summary, various pertinent points may be considered in using this approach to determine the accuracy of the various algorithms.

a. Clearly GPCP always gives results of a superior geometric accuracy, irrespective of surface type or grid size.

b. In highly autocorrelated surfaces, GINOSURF, CONSYS and indeed any 'centre-point triangulation' method of isarithm interpolation are superior, although this gradually changes to 'diagonal triangulation' (GHOST) and 'direct-grid' (SURFACE II GRAPHICS) interpolation as the

surface autocorrelation coefficient decays. While SACM was only evaluated at the full-grid size, it appears to be consistent with SURFACE II GRAPHICS, both from the coefficient and the shape of the correlation curve.

c. Interestingly, the bottom graphs for each method suggest that all interpolation methods consistently interpolate low. This may be a feature of the data sampling as discussed in 3.4.4. The grid was observed in point mode and therefore should be of a higher accuracy than the individual reference isarithm points which were observed in stream mode. It may be that when grid points were being observed, the 'floating mark' in the plotting machine was closer to ground level than when the isarithms were being observed, introducing a systematic error.

d. The maximum isarithm length and the maximum area discrepancies occur when 60-65% of the total map area is enclosed by an isarithm. It may therefore be assumed that there is a systematic error component involved in determining the accuracy utilising this method, which might lower the correlation coefficient if the majority of isarithms were long.

7.2.2.2 Position And Elevation Error

While the isarithm correlation coefficient may provide a good statistic to describe relative geometric accuracy, it does not describe the absolute geometric accuracy. This requires the position and elevation errors (Figures 7.13 to 7.15) and the Koppe accuracy formulae (see Figures 7.16 to 7.18).

Significantly, most of the data set/package combinations produce elevation error curves which are similar, scaled-down versions of the position error curves. This would suggest a consistent, balanced error/gradient relationship within each of the data sets. Notable exceptions to this are the performances of SURFACE II GRAPHICS (FORV-17, FORV-33, INCH-26 and LIAN-26), SACM (FORV-65), GHOST (LIAN-13) and CONSYS (FORV-17). From the graphs, it may be assumed that SURFACE II GRAPHICS is a particularly poor interpolator in flat areas, since elevation error is proportionately smaller than position error. This effect seems limited to medium-resolution grids and rough micro-relief surfaces. SACM (FORV-65) and GHOST (LIAN-13) suffer similarly on smooth surfaces, although CONSYS (FORV-17) has a disproportionately high amount of error in steep areas of rough micro-relief surfaces. Therefore, the coarser methods of isarithm production (direct-grid and diagonal-triangle) have most error in flat areas and the finer (centre-point) method has most error in steep areas. Overall, the graphs support the proposition that GPCP is consistently superior irrespective of resolution or surface. GINOSURF and CONSYS are generally good although GHOST is marginally

Figure 7.13 Grid-to-isarithm Lindig errors (INCH)

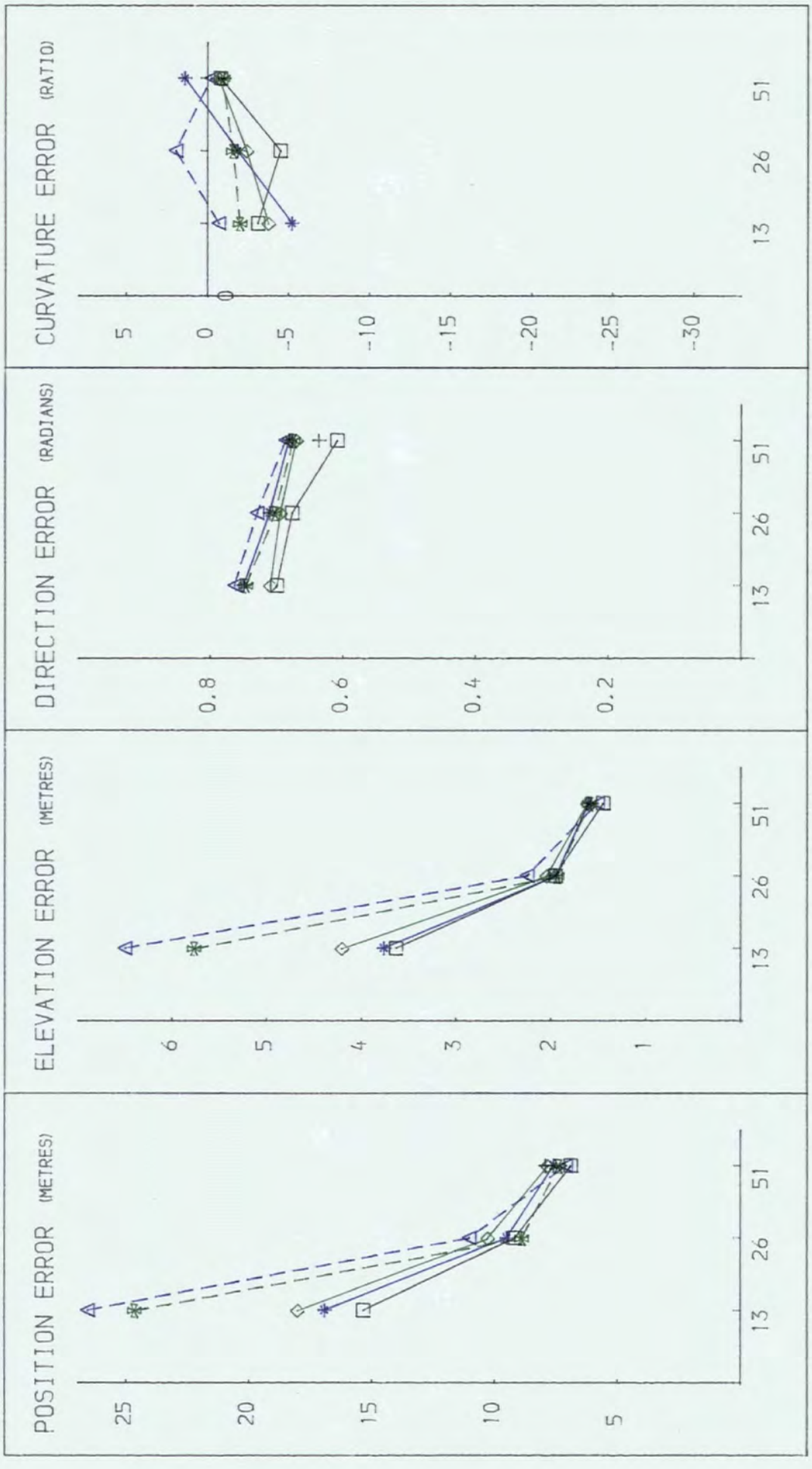


Figure 7.14 Grid-to-isarithm Lindig errors (LIAN)

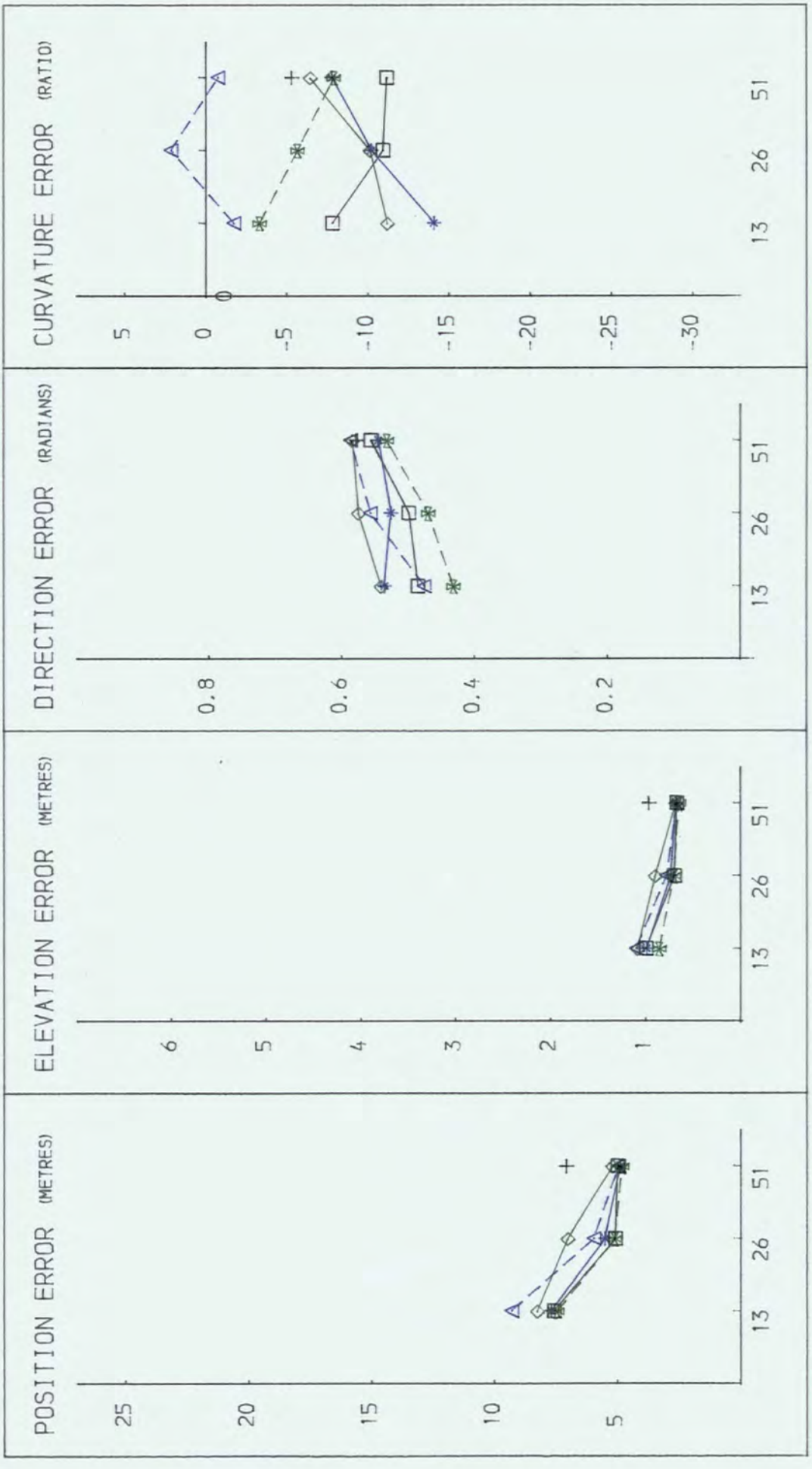
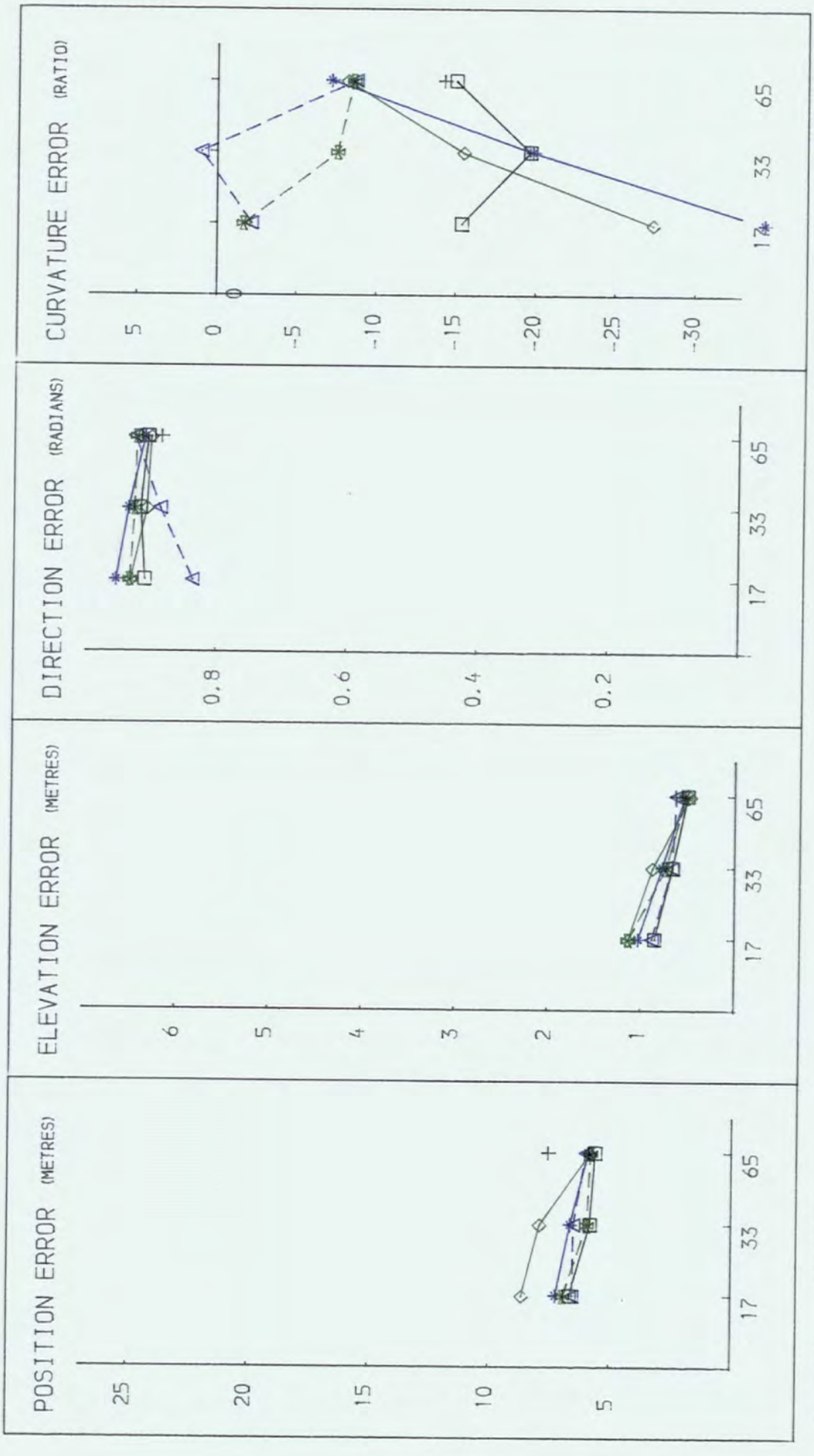
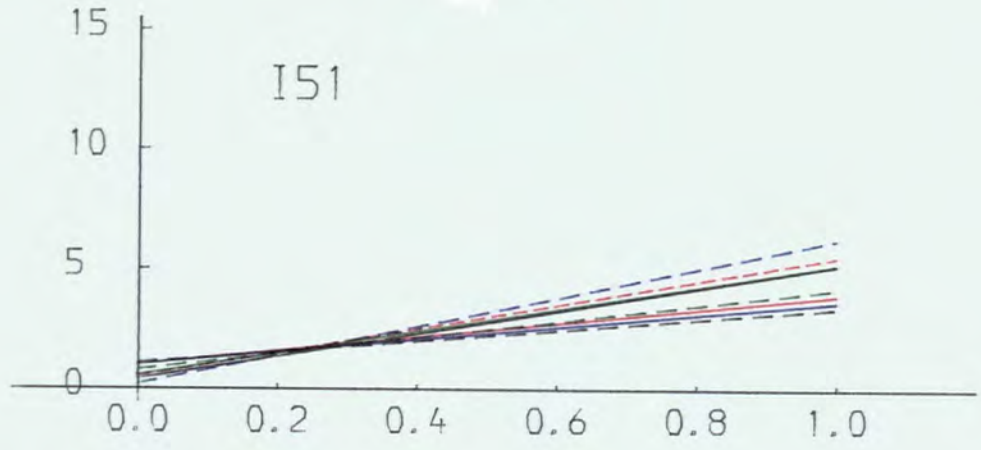
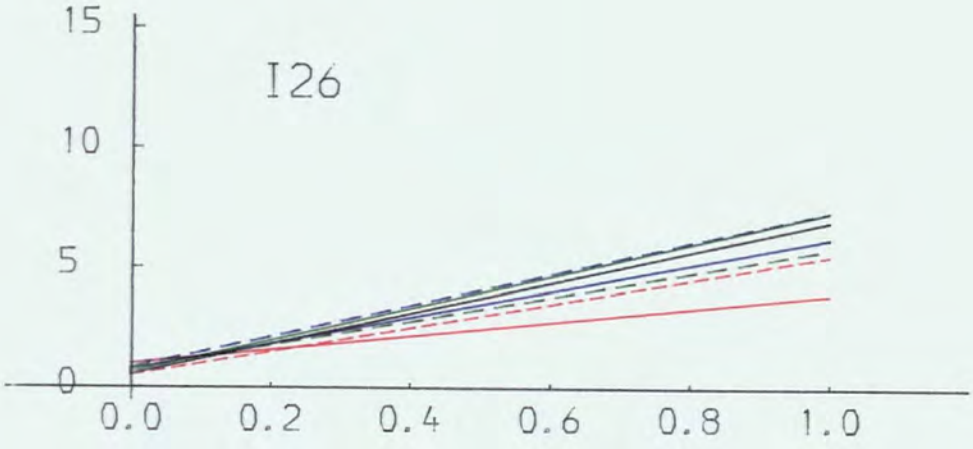


Figure 7.15 Grid-to-isarithm Lindig errors (FORV)







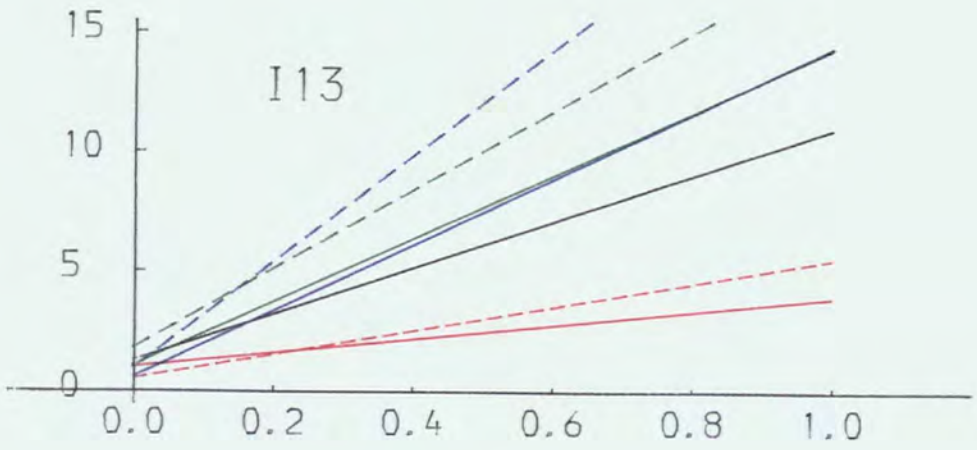
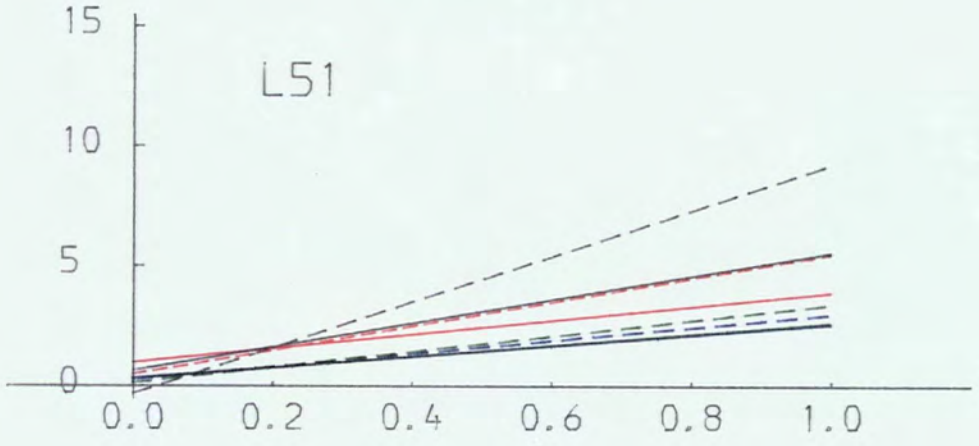
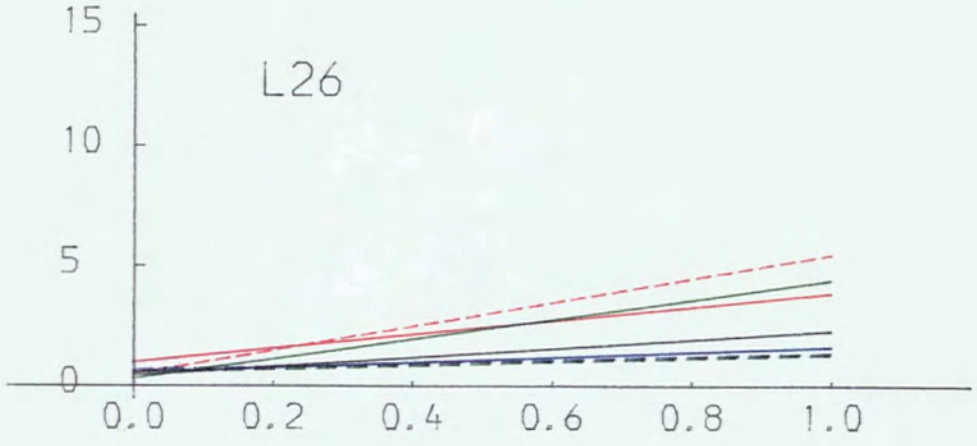


Figure 7.16 Grid-to-isarithm Koppe graphs (INCH)





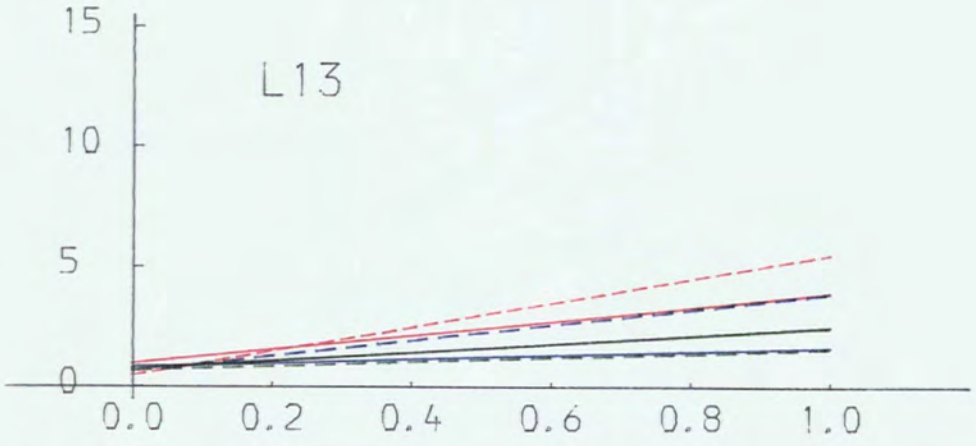
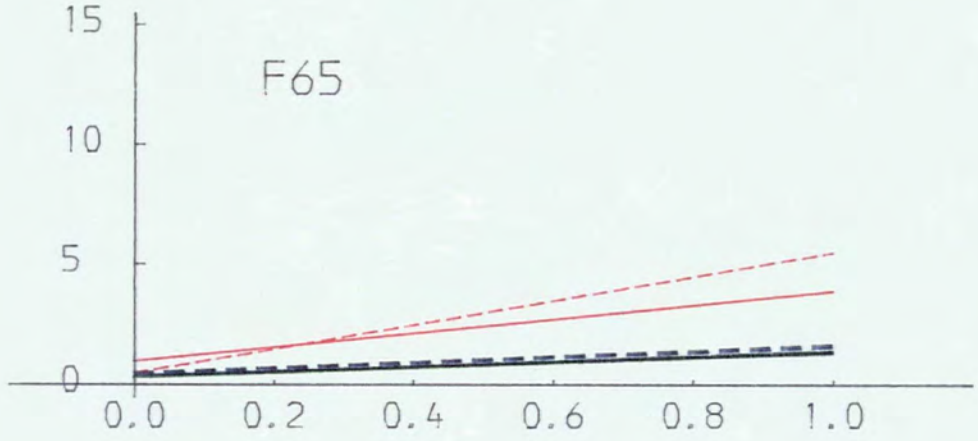
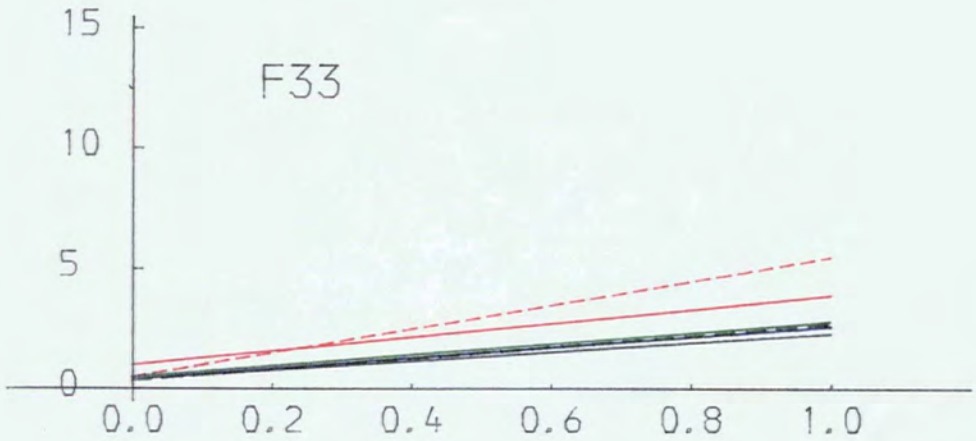


Figure 7.17 Grid-to-isarithm Koppe graphs (LIAN)





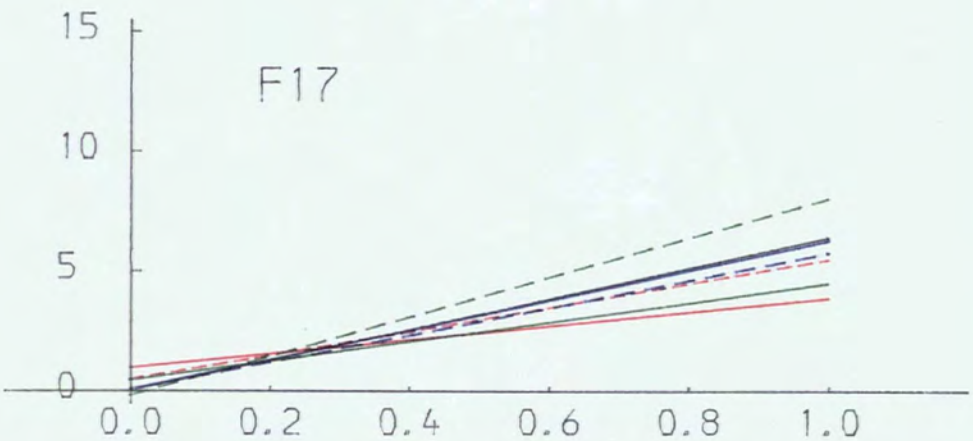


Figure 7.18 Grid-to-isarithm Koppe graphs (FORV)

superior in FORV.

The benefit of utilising this scale-dependent measure of accuracy evaluation is underlined by considering the magnitude of the errors. LIAN and FORV, which cover similar areas and have similar grid resolutions, produce reasonably consistent results irrespective of grid resolution. INCH produces an excessive variation between the 13 and 26 grid sizes - however INCH does cover a much wider range of values in all three dimensions and the respective grid resolutions are 13 (120m), 26 (60m). Even for the finest resolution (51), position (10.0m) and elevation (2.0m) errors are twice those of FORV and LIAN (6.0m and 0.8m), although this must be balanced against a similar ratio of ranges in 'z' - FORV (28m), LIAN (90m) and INCH (180m). Therefore, results using position and elevation errors to define accuracy show a remarkably high correlation with the surface area analysis (see Table 4.6), whereas isarithm correlation results are highly correlated with the scale-independent autocorrelation characteristics (see Table 4.3). Additionally, the latter, as a result of the systematic effect of line length on the descriptor, must be considered to combine some measure of morphological trueness with geometric accuracy.

The ability of the interpolated isarithms to meet required map accuracies may be examined by considering the relevant graphs summarising the Koppe formulae. In considering them, it should be noted that the overlay maps are approximately at the scales

a. INCH - 1:13,000

b. LIAN - 1:8,000

c. FORV - 1:6,000

Excluding a few individual cases and FORV (17), all FORV and LIAN maps fall within the required accuracies of the German 1:5,000 and Swiss 1:10,000 scale maps. In addition, while the INCH maps do not attain the required standard, it should be remembered that they are at a smaller scale, and therefore INCH (26 and 51) could still fall within the accuracy for a comparable scale (see Figure 5.1). The extremely poor INCH-13 illustrates the problems of interpolating from a grid with a coarse resolution in relation to the features being mapped.

Unfortunately, as the variance across data sets is too extensive, no meaningful pattern emerges from the graphs with respect to the performances of specific packages. This emphasises the importance of the descriptor purely as a binary statement of accuracy - either the product is within or not within the allowable tolerance.

7.2.3 Morphological Trueness

Two expressions for morphological trueness of the interpolated isarithms were calculated using CONLIN. The resultant errors (direction and curvature) are summarised with the position and elevation errors, in Figures 7.13 to 7.15.

Morphological trueness is related to the accuracy of the relative shapes and orientations of the isarithm under interest wth regard to the reference isarithm - no consideration of the relative positions of the two isarithms is made.

Direction error expresses the error between the directions of the line segments on the interpolated isarithm with comparable points on the reference isarithm. If the isarithms are parallel, the error will be Oradians. Any deviation from parallelism results in an error which, for the purpose of this evaluation, has been reduced to its absolute magnitude. The shape of two isarithms may be identical, but, if the orientation is wrong then direction error will exist, although there will be little curvature error.

Curvature error expresses the ratio between the radii of curvature of the interpolated isarithm and the reference isarithm. Section 5.2.2 and Figure 5.2 show that the magnitude of the error varies with R, which is established a priori and was set to 500.0m throughout this examination. Thus, whereas the other three graphs in Figures 7.13 to 7.15 show absolute error, the magnitude of the curvature error only reflects relative error. However, the sign of the error is important. A positive error implies that the radius of curvature of the interpolated isarithm is greater than the radius of curvature of the reference isarithm. Alternatively, a negative error implies that the radius of curvature of the reference isarithm is greater than that of the interpolated isarithm. Clearly the smaller the magnitude of the error, the greater the similarity of the shape of the isarithms.

7.2.3.1 Direction Error

An important aspect of the graphs of direction error is the low variance within any data set across both resolution and

interpolation package. However, a strong correlation exists between error magnitude and the surface autocorrelation of the differing data sets.

LIAN shows the most error variance of all the surfaces. However, since the surface is a highly autocorrelated linear one, the magnitude of the error is considerably lower than that from the other surfaces (approximately 0.5radians). Importantly, coarser resolution grids produce lower errors than finer resolution grids. Since the isarithms of the coarser resolution grids only portray macro-features - which in the case of this surface are smooth and linear - direction of both interpolated and reference isarithms will be near similar. At the finer resolution, the rougher micro-features are detected although they are not delineated so well - hence the poorer results. In general, the packages which involve generating finer resolution intermediate grids are superior, thus the order of increasing error is CONSYS, GPCP, GINOSURF, GHOST, SURFACE II GRAPHICS and SACM. Where smooth surfaces occur, the initial pre-threading interpolation to increase the grid resolution before isarithm threading (for example CONSYS and GPCP) produces high accuracy results, providing a more rigorous basis for isarithm threading.

Within INCH, the error variance is much lower although the average magnitude is higher. INCH is a rougher, less autocorrelated surface than LIAN with a more complex macro- and micro-structure. Thus the coarser resolution grid produces generalised isarithms which result in high direction error. As resolution increases, isarithm representation of

macro-features increases and therefore error decays. As suggested previously, the ability of the grid-to-triangulation interpolation process effectively to increase the resolution of the grid decreases as the surface roughness increases. Therefore, the triangulation-based methods (CONSYS, GINOSURF and GHOST) produce inferior products to the function and grid methods (GPCP, SACM and SURFACE II GRAPHICS).

FORV is a micro-featured surface of low autocorrelation. Therefore error is high at the coarsest resolution since the basis for even a reasonable macro-feature delineation does not exist due to the micro-dominant nature of the surface. Indeed, with the exclusion of GHOST, the ability of the interpolated isarithms to replicate the reference isarithms only marginally improves with resolution. The nature of the surface causes problems with the grid-to-triangulation methods which require pre-threading interpolation. Therefore, the methods which involve no pre-threading interpolation (SACM, GHOST and SURFACE II GRAPHICS) are superior to those requiring interpolation (GPCP, CONSYS and GINOSURF).

In general, a feature of this form of error is the superior ability of GPCP to replicate the reference isarithms. While GPCP may not always be the optimum method it consistently produces a higher accuracy product.

7.2.3.2 Curvature Error

The curvature error shows the most consistency of all the Lindig errors between the models. While magnitude may vary quite distinctly (as discussed earlier), the relative shapes

of the graphs between each model are virtually identical. The packages that do not use smoothing produce the smallest error (GHOST and CONSYS) although this trend, especially in FORV, is reversed at the full-grid. This suggests that smoothing is not effective with coarser resolution grids - interpolated isarithms being too rounded. However, as resolution increases, the magnitude of the error as a result of post-interpolation smoothing diminishes (quite dramatically for FORV). GPCP gives consistently poor results which illustrate its ability to provide an unbiassed interpolator as regards grid resolution.

Undertaking a cross-model evaluation is more difficult. The INCH surface, while made of a larger macro-feature and thus covering a much larger area than the other surfaces, has a relatively curved terrain. Radii of curvature are therefore consistently sized at an average length. Since the interpolated isarithms can match the macro-features sufficiently well, curvature error is thus minimised. Alternatively, the LIAN surface, which covers a smaller area, has predominantly lineated isarithms with large radii of curvature. With the superimposition of micro-features with much smaller radii of curvature, any small mis-match of interpolated and reference isarithms will have a maximum effect on curvature error - hence the error magnitude is larger than for INCH. FORV isarithms contain such a variety of radii of curvature, due to the micro-nature of the surface, that the curvature error may be affected to the extreme. At the coarse grid resolution where only macro-features are delineated, the error is much greater than at the full-grid when micro detail is delineated.

7.2.4 Conclusion On Grid-to-isarithm Methods

It should be apparent from the previous discussion that, while the grid-to-isarithm process does not produce such a large range of accuracies as does random-to-grid interpolation, nevertheless accuracies may be variable. It has been shown that, excluding INCH-13 and taking into account data point error as discussed in Chapter 3, accuracies using the grid-to-isarithm programs are comparable with national mapping accuracy standards.

The various forms of accuracy may be directly associated with the surface characteristics of the data under investigation.

a. Morphological trueness (as expressed by direction error) and geometric error (as expressed by the scale-independent resemblance correlation) are strongly correlated with surface autocorrelation.

b. Morphological trueness (as expressed by curvature error) is strongly correlated with vector dispersion analysis.

c. Geometric accuracy (as expressed by position and elevation errors) is strongly correlated with surface area analysis.

The relative accuracies which may be achieved by specific

packages must be considered in conjunction with the computing resources necessary for the packages' operation. Table 7.2 summarises the cpu time (seconds) used and the virtual memory (pages per minute where one page equals 4096 bytes) required for each run of FORV data. Table 7.3 summarises the same statistics in terms of the commercial charging rate at NUMAC (in pence). This does not include costs attributed to file storage, printer and plotter output, data I/O operations, data editing or elapsed time. All statistics are sub-divided into the main program run and any other associated runs (i.e. GPCP - reformating data; GHOST, CONSYS and GINOSURF - compiling program). Comparable values for SURFACE II GRAPHICS and SACM have been estimated.

It may be seen that while GPCP interpolates isarithms of a superior and consistent accuracy, the total cost is considerably greater than by any other method - both in terms of CPU time and memory requirements. This is partly a result of each pre-threading patchwise interpolation being repeated for each isarithm that passes through the relevant patch. 'Diagonal' (GHOST) and 'direct grid' (SURFACE II GRAPHICS and SACM) methods involve similar charges and as a general 'rule-of-thumb', produce good results with data from rough surfaces. 'Centre-point' methods (CONSYS and GINOSURF) are preferred with smooth surface data.

As regards isarithm smoothing, few differences exist between for instance, CONSYS and GINOSURF. At finer resolutions, isarithm smoothing is unnecessary while at coarser resolutions it may often prove inconvenient,

		Main Run		Associated		Total	
		Memory	Time	Memory	Time	Memory	Time
CONSYS	65	1.95	2.98	0.82	0.91	2.77	3.89
	33	0.80	1.64	0.82	0.91	1.62	2.55
	17	0.47	1.21	0.82	0.91	1.29	2.12
GHOST	65	2.38	3.47	0.72	0.85	3.10	4.32
	33	0.90	1.73	0.72	0.85	1.62	2.58
	17	0.53	1.28	0.72	0.85	1.25	2.13
GINOSURF	65	4.63	4.99	0.67	0.83	5.30	5.82
	33	2.12	2.77	0.67	0.83	2.79	3.70
	17	1.47	2.16	0.67	0.83	2.14	2.99
GPCP	65	11.07	7.33	2.83	3.91	13.90	11.24
	33	4.47	3.32	0.9	1.62	5.37	4.94
	17	2.30	2.03	0.5	1.13	2.35	3.16
SACM	65	4.50	4.50			4.50	4.50
SURFACE II	65	4.50	4.50			4.50	4.50
	33	2.00	2.60			2.00	2.60
	17	1.30	2.00			1.30	2.00
DTC	FB2	16.30	5.03	9.12	5.07	25.42	10.10
	FC	16.52	5.10	9.12	5.07	25.64	10.17
	FS	16.50	5.08	9.12	5.07	25.62	10.15

Table 7.2 CPU time and virtual memory requirements for isarithm generation (in page/minute and seconds)

		Main Run		Associated		Total
		Memory	Time	Memory	Time	Cost
CONSYS	65	1.8	80.5	2.4	24.6	109
	33	1.3	44.3	2.4	24.6	73
	17	1.1	32.7	2.4	24.6	61
GHOST	65	2.6	93.7	2.3	23.0	122
	33	1.4	46.7	2.3	23.0	73
	17	1.1	34.6	2.3	23.0	51
GINOSURF	65	2.5	134.7	2.2	22.4	162
	33	2.1	74.8	2.2	22.4	102
	17	1.8	58.3	2.2	22.4	85
GPCP	65	4.1	197.9	2.0	105.6	310
	33	3.6	89.6	1.5	43.7	138
	17	3.1	54.8	1.2	30.5	90
SACM	65	2.7	121.5			124
SURFACE II	65	2.7	121.5			124
	33	2.1	70.2			72
	17	1.8	48.6			50
DTC	FB2	8.8	135.8	4.9	136.9	286
	FC	8.7	137.7	4.9	136.9	288
	FS	8.8	137.2	4.9	136.9	288

Table 7.3 CPU time and virtual memory costs for isarithm generation (in pence)

generating cusps and overlapping isarithms. The extra cost of using such a technique (approximately 50% more) is only to be encouraged in medium resolution grids.

Alternatively, cost may be greatly reduced by utilising a smaller grid. Cost decays exponentially with decreasing grid size, and therefore a grid size reduction when considering highly autocorrelated surfaces, e.g. LIAN, and to a lesser extent INCH, is beneficial and would allow a superior package to be utilised for the same cost.

Therefore, in considering the optimum grid size required before grid-to-isarithm interpolation is attempted, the user must consider several points. He must consider;

a. which algorithm will be used and whether it includes smoothing.

b. the smoothness of the surface - a smooth surface may allow a coarser grid resolution.

c, the nature of the surface - whether it contains micro-features which may be lost if too coarse a grid resolution is used. The grid resolution should be half the smallest wavelength of a feature.

d. the cost and storage limitations both in random-to-grid and grid-to-isarithm interpolation - both

increase dramatically with increasing grid resolution.

7.3 RANDOM-TO-ISARITHM

Random-to-isarithm interpolation was studied using DTC to generate isarithms of the various point data sets. Unfortunately, as discussed in A1.4.6, the 'GRID' point data sets could not be used with this program. Only eleven data sets were therefore processed using DTC. Throughout, no smoothing of isarithms was applied.

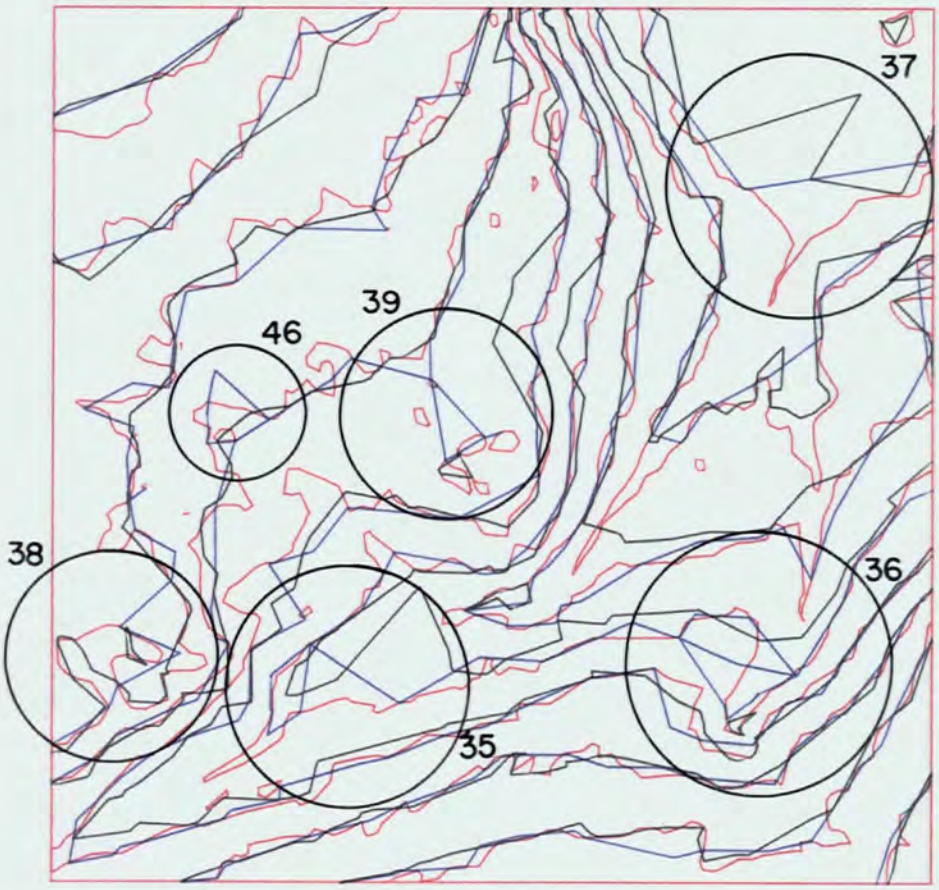
7.3.1 Graphical Evaluation

Three overlay isarithm maps were prepared (see Figures 7.19 to 7.21) using the colour and line types as specified in Table 7.1b.

In general, the macro-features of all three models are presented at the expense of over-generalised micro-features. In INCH, excluding some ambiguities (35 and 36) with CONTOUR data, all data sets generate isarithms of reasonable quality. SCATTER is particularly successful (36). The rest of the map area is adequately replicated with the exceptions of BREAKLINE (37), the corrie (38) - which is traced in four completely different ways - and RIVER data, which as expected, over-generalise away from the rivers. Micro-features are virtually non-existent with the exception of four different interpretations of the surface at 39.

As with grid-to-isarithm interpolation, LIAN data perform

Figure 7.19



INCH : POINT DATA

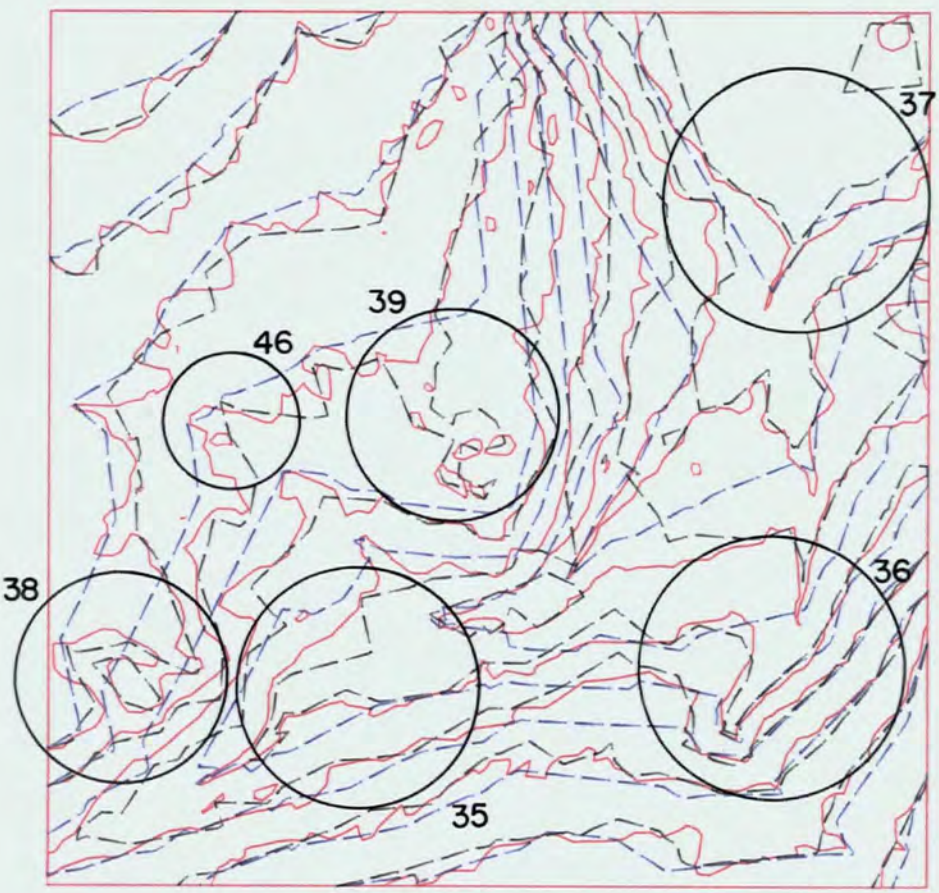
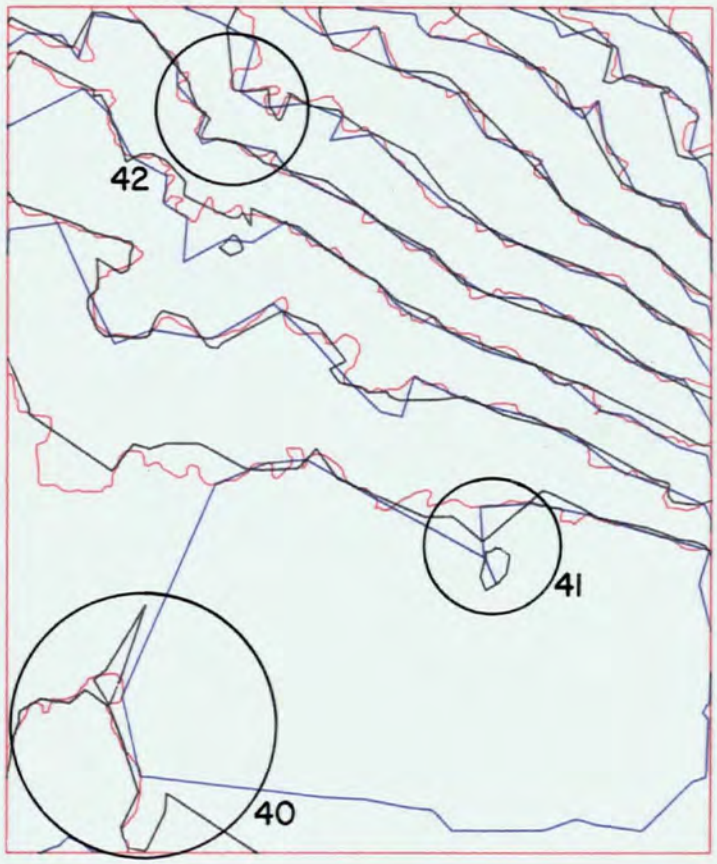


Figure 7.20



LIAN : POINT DATA

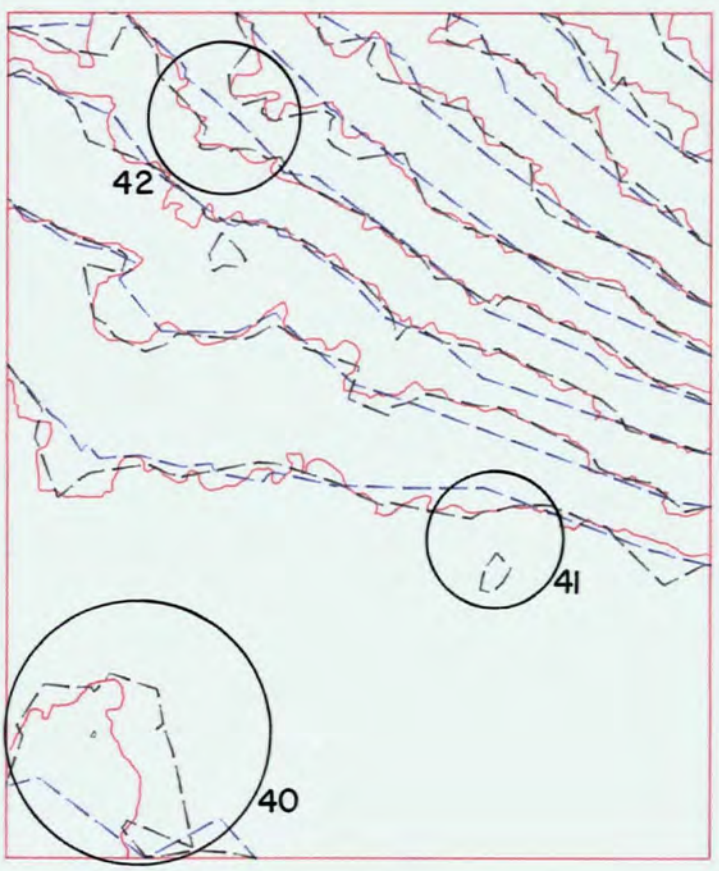
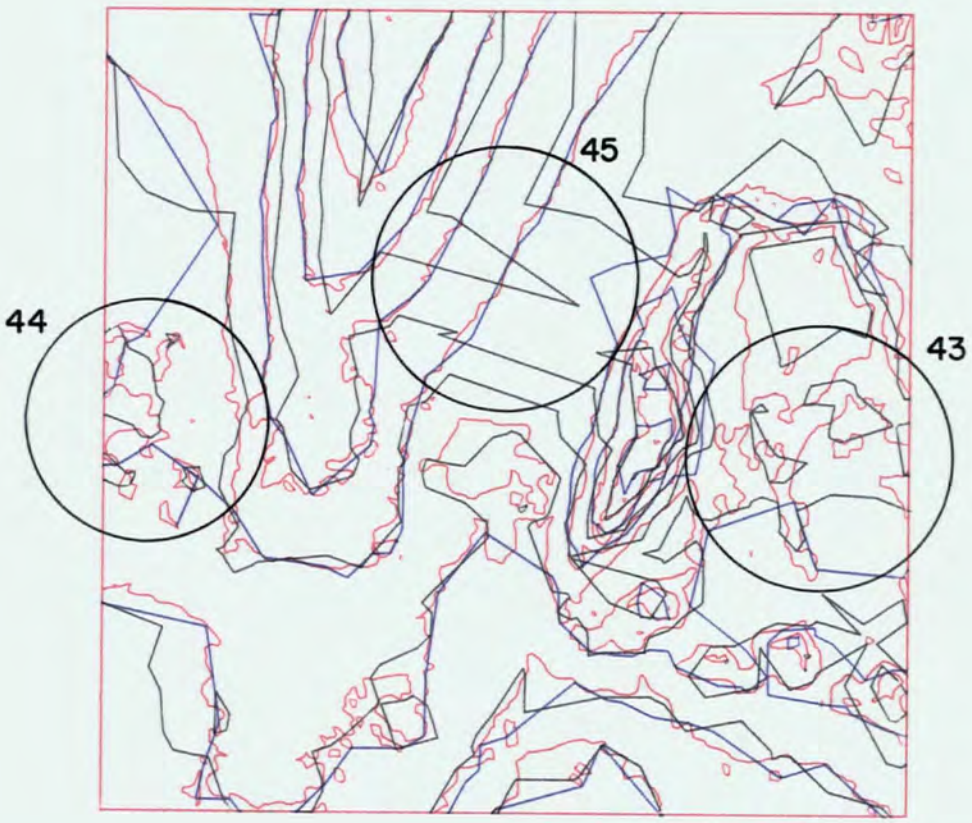
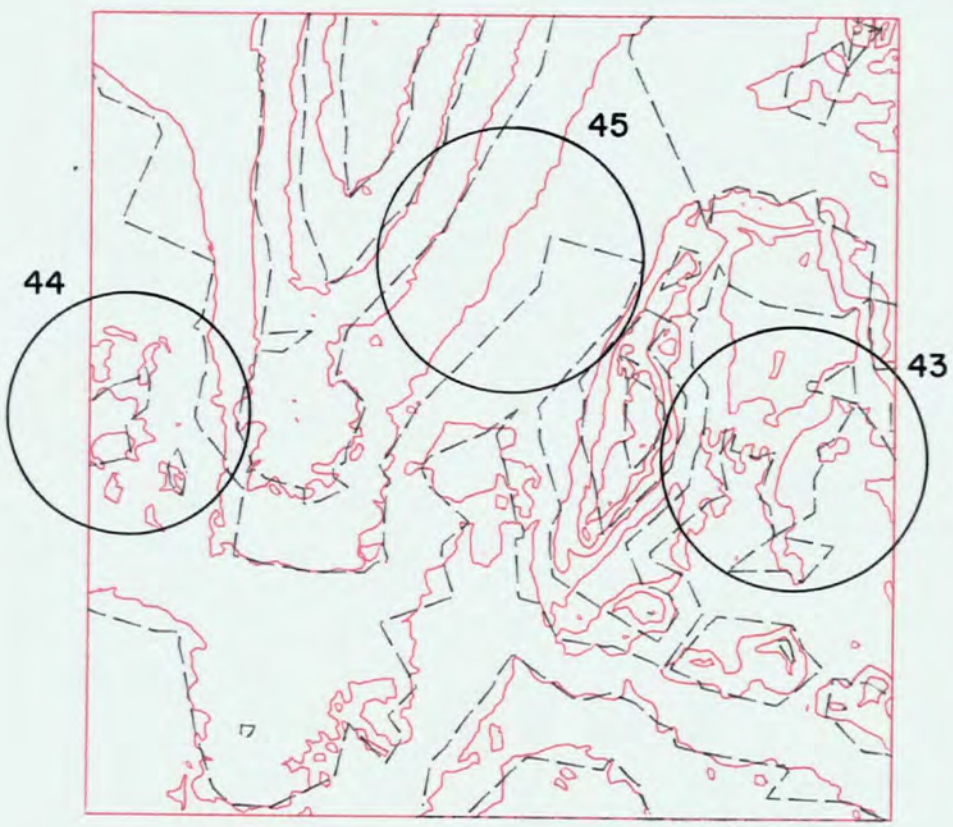


Figure 7.21



FORV : POINT DATA



adequately well. While the main highly autocorrelated macro-features are easily replicated, problems occur with the micro-features (e.g. 40, 41 and 42) and only BREAKLINE and SCATTER data perform sufficiently well in these places.

Within FORV, SCATTER data are generally successful for both macro- and micro-relief (e.g. 43); CONTOUR data, while successfully defining macro-relief, have distinct problems with micro-relief (43 and 44), and BREAKLINE data have problems with both aspects (43 and 45).

As a result of the size of the data sets involved (approximately 250 points only within the map borders), the isarithms in general tend to be angular. RIVER data are especially so in both LIAN and INCH since most points are confined to very limited areas (see Figure 3.4). However, angularity and linearity cannot be directly correlated with the NNS, as CONTOUR data in LIAN (1.340) and INCH (1.251) appear worse in this respect than BREAKLINE data (LIAN - 1.050, INCH - 1.094).

Ambiguities represent a serious problem, especially with linearly-derived data. SCATTER and RIVER data perform satisfactorily, however CONTOUR and BREAKLINE data produce distinct associated effects. Within CONTOUR data, the 'T' effect is visible (INCH-39, LIAN-41 and FORV-44). This is a result of the occurrence of a series of points with a 'z' value equivalent to the isarithm being generated. Threading is satisfactory until another point of similar 'z' value is discovered off the main line. The isarithm is then threaded out to the point (technically in one triangle) and back

(technically in the contiguous triangle). Unfortunately, since the two common vertices have an equivalent 'z' to the isarithm level, this results in a 'T'. A 'DOUBLE LOOP' effect (INCH-36 and -46) is a mutation of this, where at least two vertices are included in the subsidiary loop. The isarithm may not retrace its exact path, but at some stage it will pass through a point which has already been visited. Both effects are a direct result of threading isarithms through a series of points with similar 'z' values.

A 'FIGURE-OF-EIGHT' effect occurs frequently in BREAKLINE data (FORV-43, LIAN-40) although it may occur in CONTOUR data (INCH-35). This is the result of a vertex (which has an equivalent 'z' value to the isarithm level) being involved in two unrelated isarithms which both pass through the same vertex.

Examining finally the possibilities of replicating the reference isarithms using the various data sets, it is apparent that SCATTER data consistently delineate both macro- and micro-relief with the required smoothness and no ambiguities. CONTOUR and BREAKLINE data produce generally acceptable products, although there are occasional gross blunders. RIVER data lead to severe angularity away from valleys. In addition, the valleys, even though the points are clustered within them, are not properly defined. The large, within-data set discrepancies are clearly visible from a cursory glance at the overlay maps and suggest the unsatisfactory nature of this technique given such a highly data-biassed result.

7.3.2 Geometric Accuracy

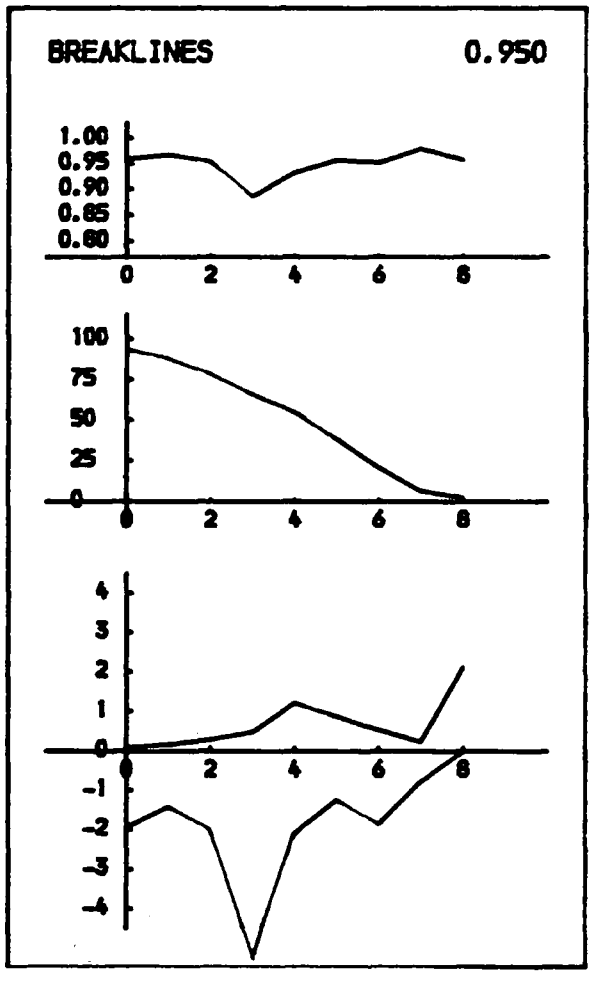
A detailed examination of the various methods of establishing the geometric accuracy of the isarithms was performed similar to the grid-to-isarithm study. The results are summarised (for correlations) in Figures 7.22 to 7.24; (for position and elevation errors) in Figure 7.25 and (for Koppe formulae) in Figure 7.26.

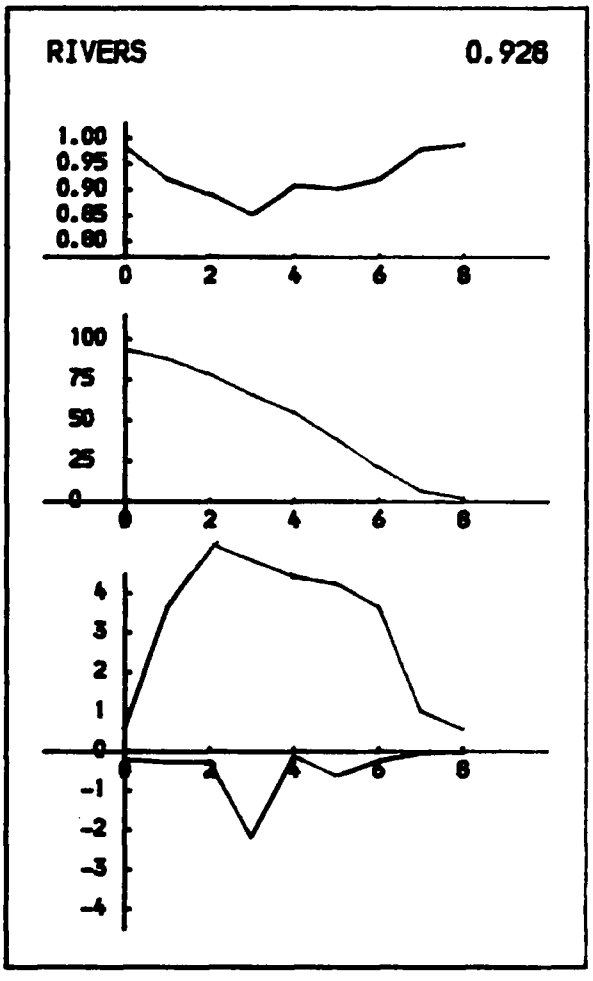
7 3.2.1 Correlation Statistics

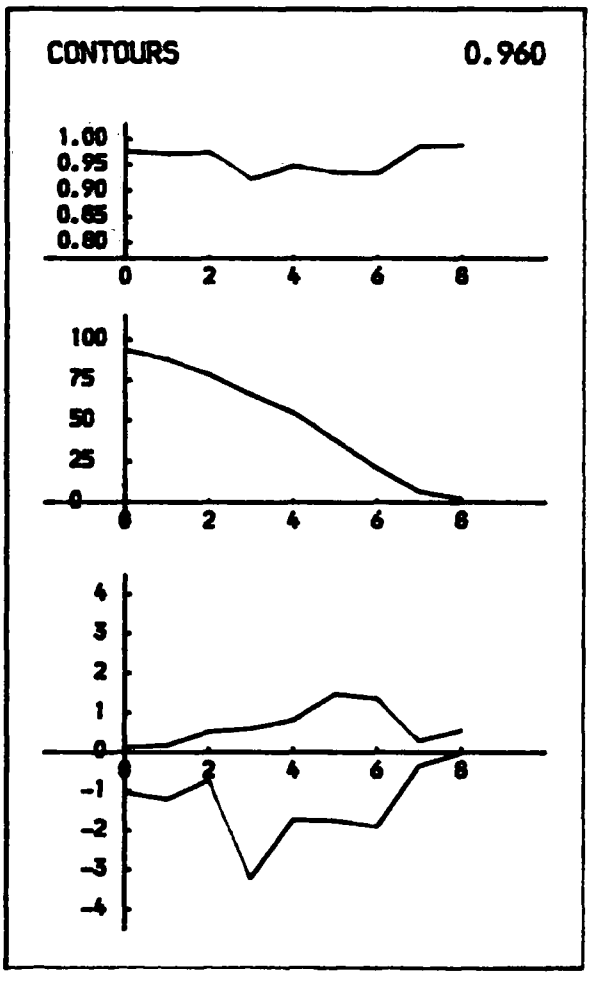
The isarithm correlation summary graphs support previously outlined theories concerning this method of accuracy evaluation and the likely interpolation accuracies of the data sets. The importance of the isarithms containing 60 - 65% of the area is underlined in all three figures, with maximum error occurring at these levels. There is similarly a strong correlation between the isarithm correlation coefficients and the summary data autocorrelation statistics as outlined in 'VECT' in Table 4.3. Additionally, within each model, there is a strong correlation between NNS - one of the powerful random-to-grid predictors - and isarithm correlation.

With relevance to the previous section, a unique signal for each data type is visible in the lowest graphs describing the areas inside the reference, yet outside the interpolated isarithms and vice versa. In general, use of BREAKLINE and CONTOUR data always leads to interpolated isarithms which are lower and therefore outside the reference isarithms. Conversely, use of RIVER data consistently leads to interpolated isarithms which are too high. Use of SCATTER data

Figure 7.22 Random-to-isarithm correlation statistics (INCH)







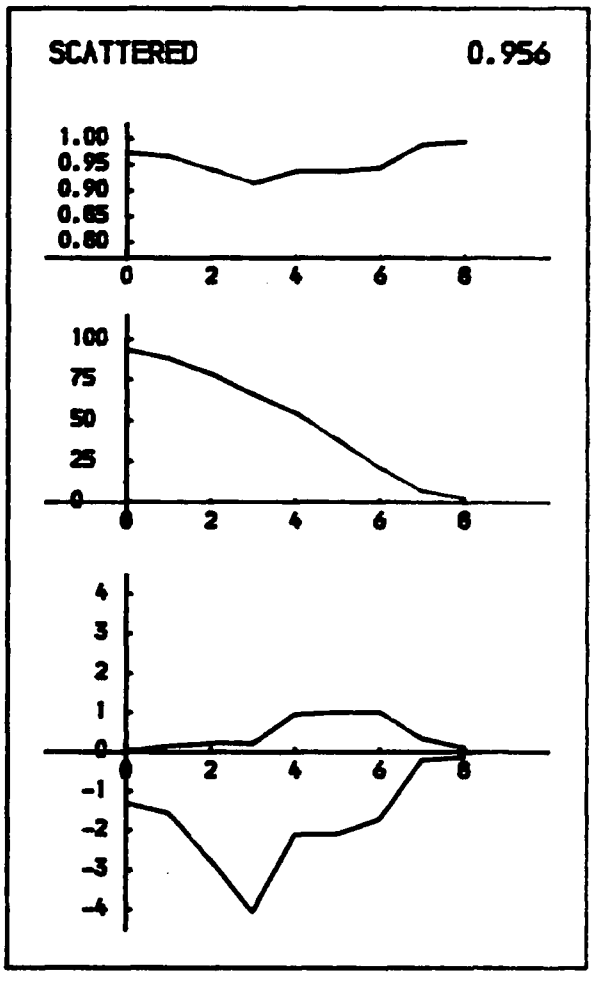


Figure 7.23 Random-to-isarithm correlation statistics (LIAN)

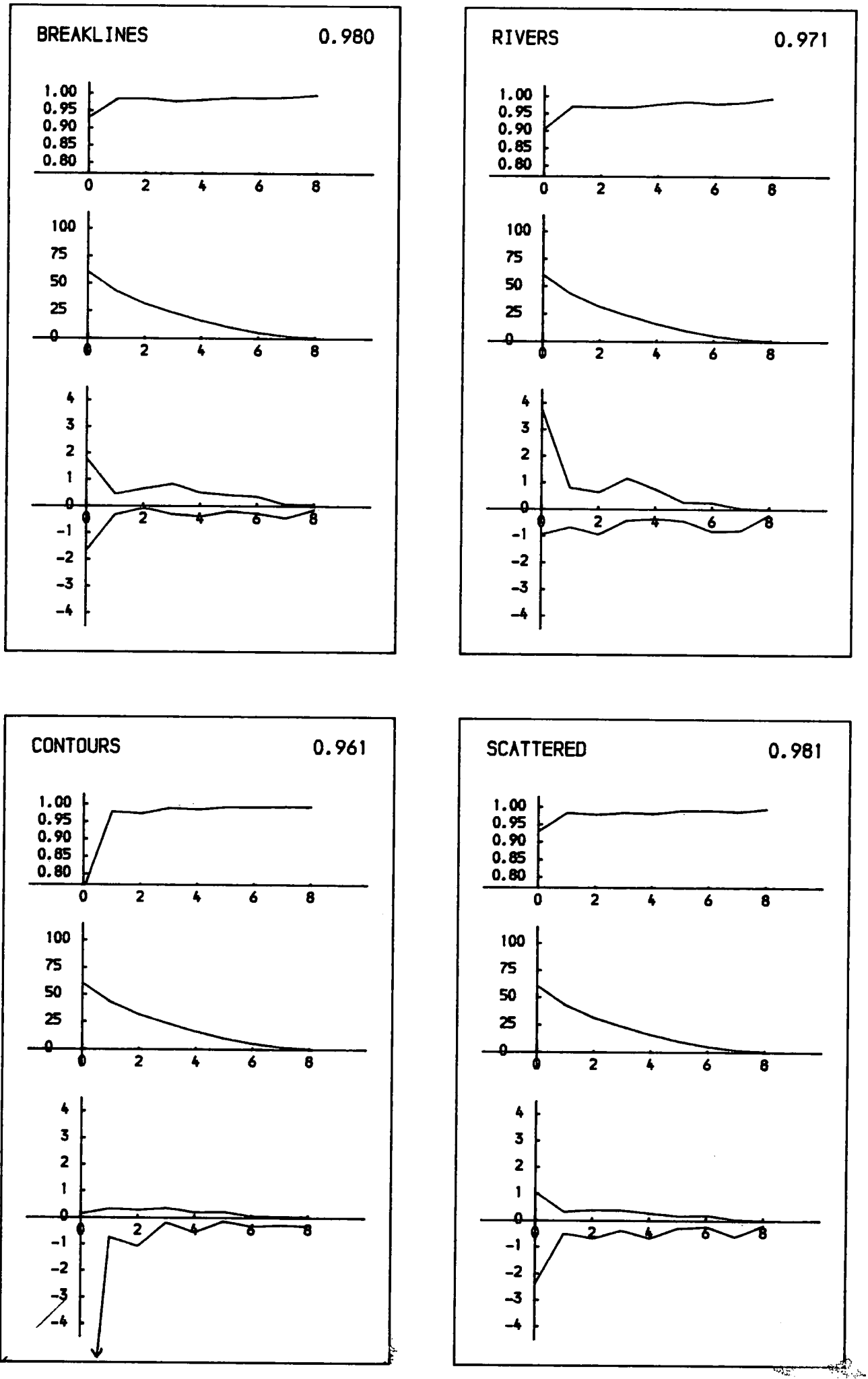
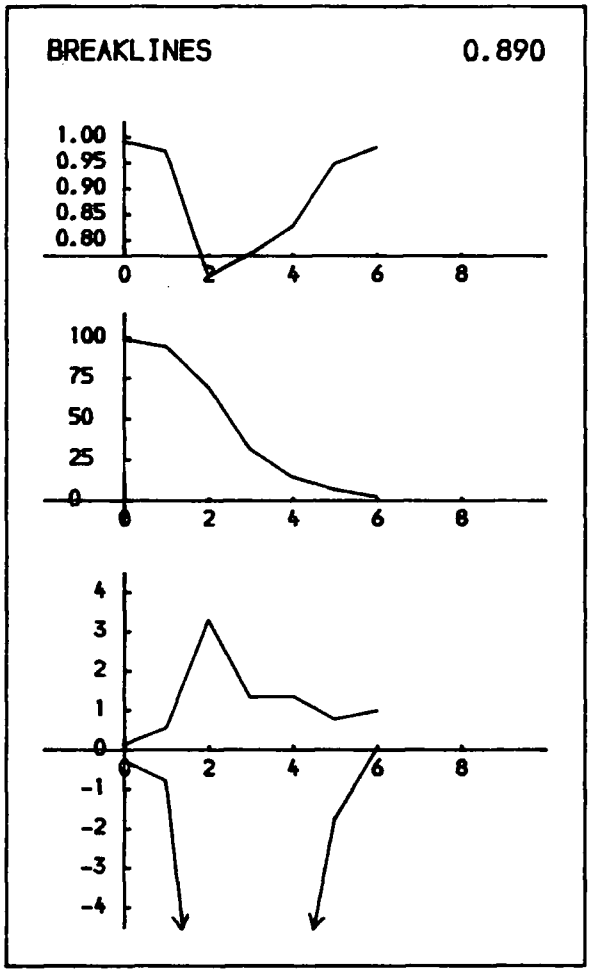
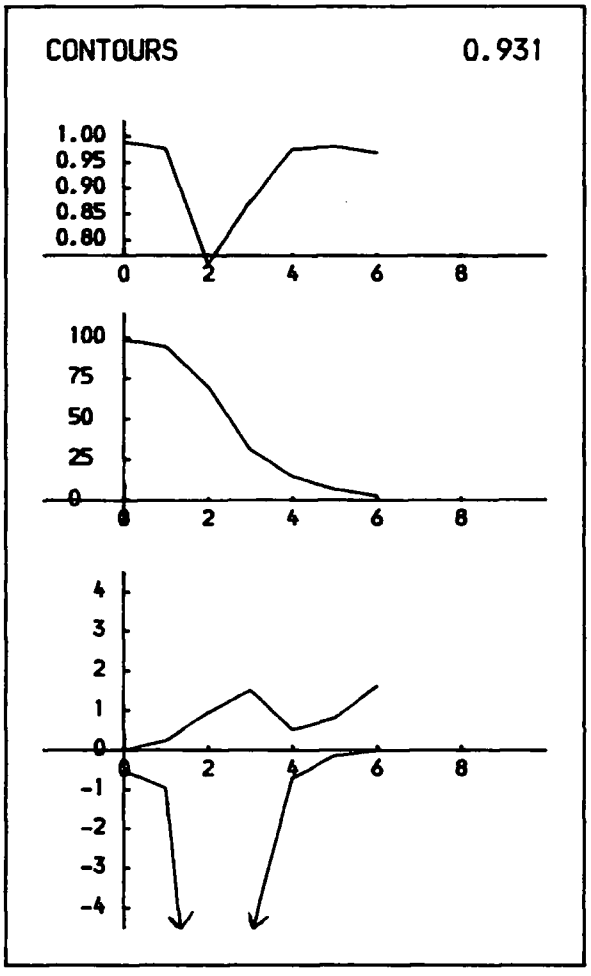


Figure 7.24 Random-to-isarithm correlation statistics (FORV)





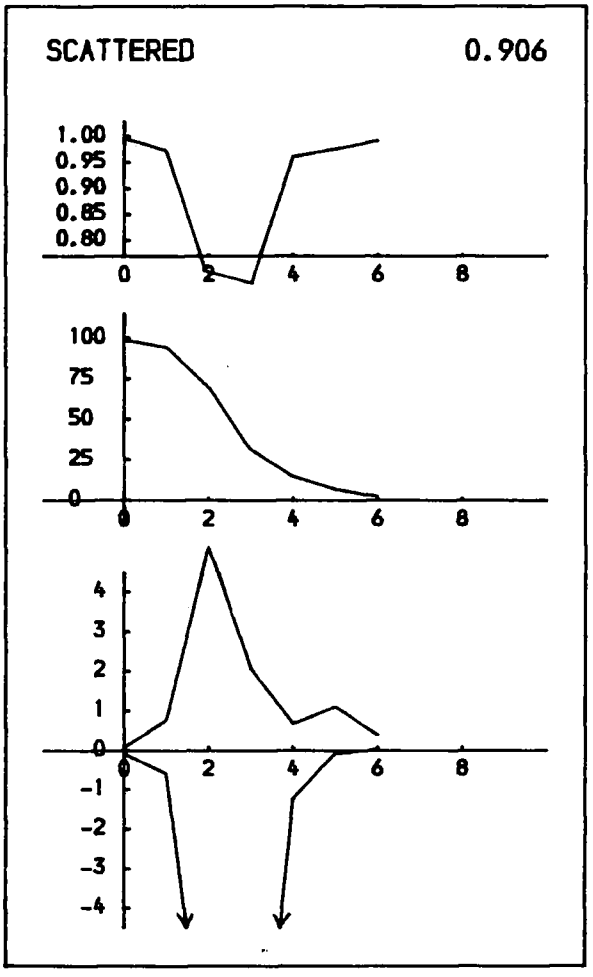


Figure 7.25 Random-to-isarithm Lindig errors

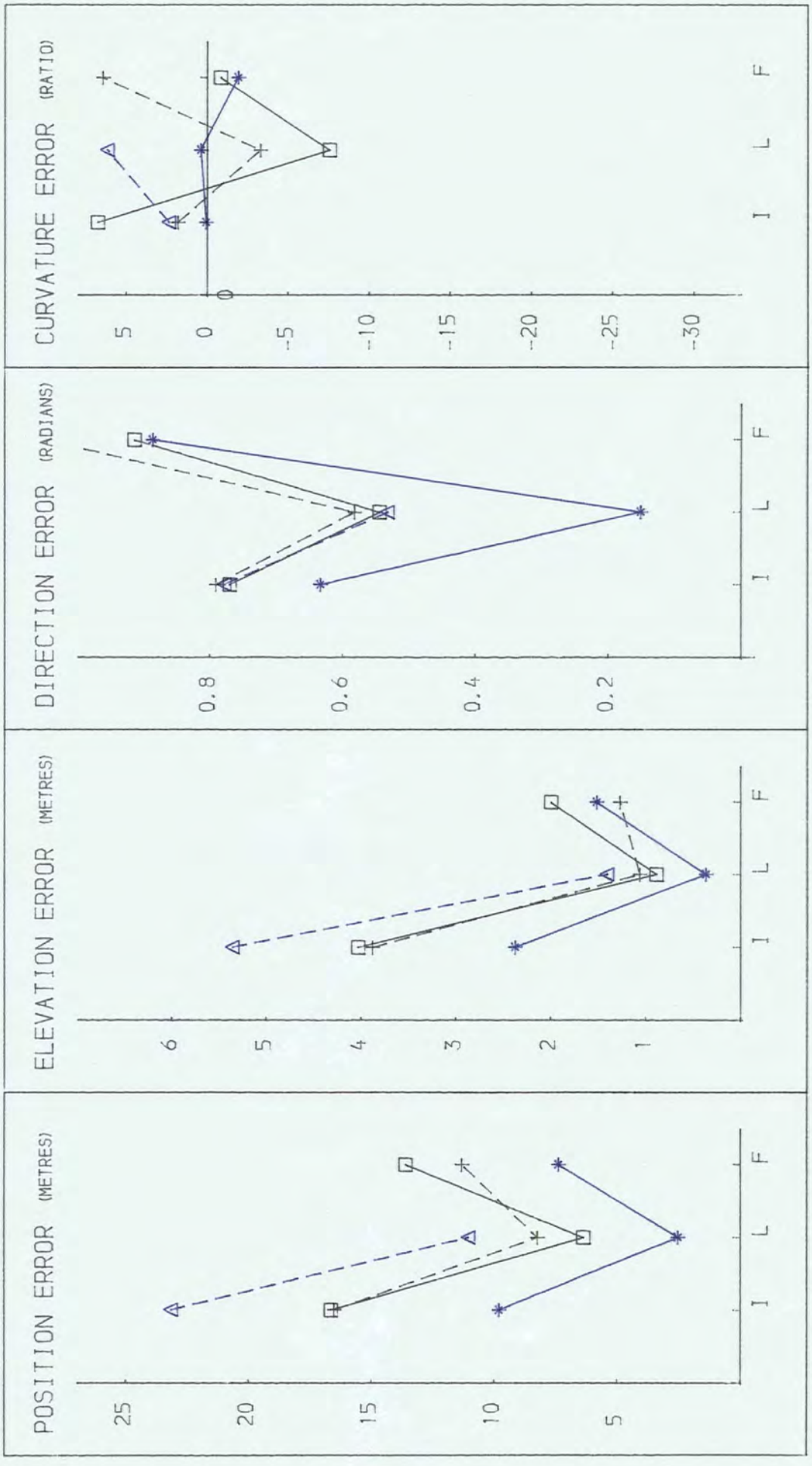
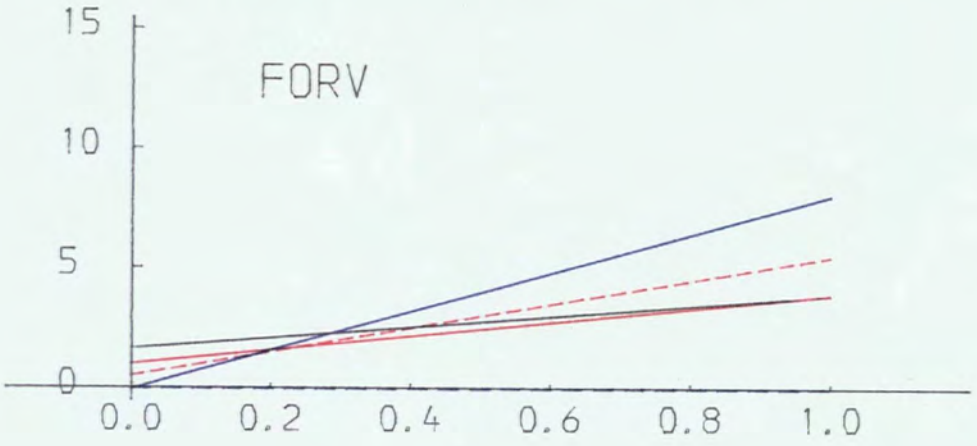
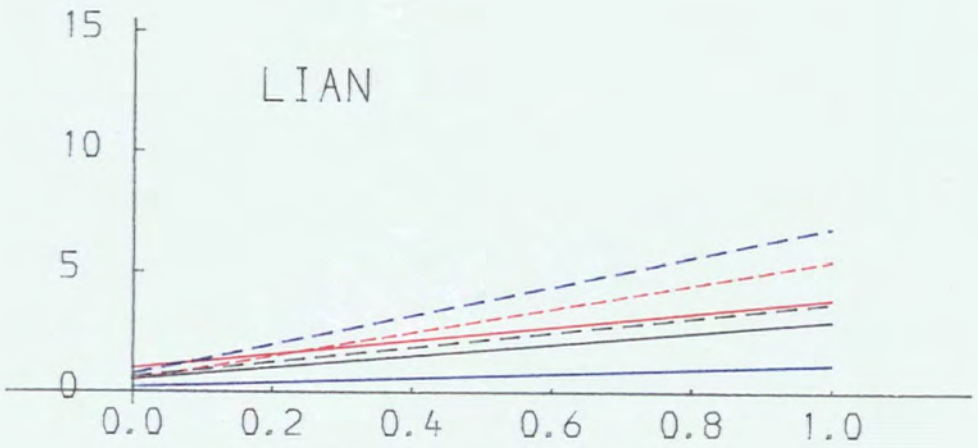
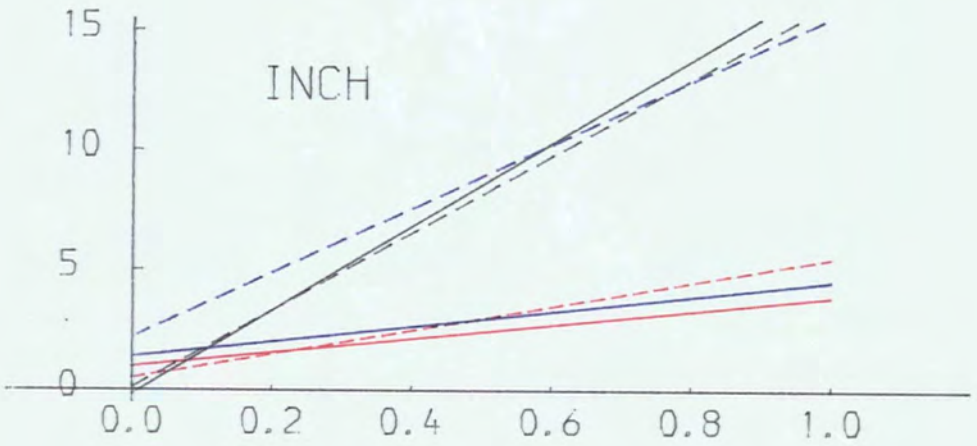


Figure 7.26 Random-to-isarithm Koppe graphs







leads to a more balanced distribution of error, although the isarithms err marginally on the low side.

7.3.2.2 Position And Elevation Error

As in grid-to-isarithm interpolation, there is a strong similarity between the position and elevation errors. This suggests that error is similarly distributed across varying gradients with the exception of FORV-SCATTER, where the majority of error appears to be on low-gradient slopes. In comparison with similar grid-to-isarithm statistics, use of all data sets leads to comparable results, although use of CONTOUR data, notably in LIAN, produces superior results. This is to be expected, since many of the data points and thus final isarithm points coincide with points on the reference isarithm with which they are being compared. Overall, a strong correlation exists between position and elevation errors and the coefficient of isarithm correlation.

With regard to the Koppe formulae, use of point data gives rise to results similar to those from the coarsest resolution grids. While use of LIAN data (excluding RIVER) produces satisfactory results, those from the use of most FORV and INCH data do not fall within the standards specified. No distinct pattern exists, with even FORV-CONTOUR falling well outside the limit.

7.3.3 Morphological Trueness

7.3.3.1 Direction Error

As in the use of the gridded data, the direction error for any model type has a low variance. The main determining factor as regards direction error is also model type. With the exclusion of CONTOUR data, there is a low within-model variance of error, with mean values of 0.78 (INCH), 0.58 (LIAN) and 0.90 (FORV), which are comparable to the results produced from the gridded data.

Use of CONTOUR data produces more variable results. With LIAN data, the majority of points which define the interpolated isarithms are those which were sampled from the reference isarithms - the error should therefore be low. INCH and FORV data contain much longer and more varied reference isarithms and therefore the interpolated isarithms do not simply trace through the points that were sampled from the reference isarithms. Hence, production of the INCH- and FORV-CONTOUR isarithms has involved more between-data point interpolation to generate the isarithms - thus resulting in direction error closer to that computed from the other data sets, but marginally lower.

7.3.3.2 Curvature Error

Curvature error in these examples is more difficult to explain. Use of CONTOUR data leads to very low error as a

result of the high correlation between data points and points on the reference isarithms. Noticeably, this is worst for FORV which is likely to contain the most points which are not data points within the interpolated isarithms. The errors arising from the use of other data sets are more variable and do not appear to follow any fixed pattern.

7.3.4 Conclusion On Random-to-isarithm Procedures

In general, the various accuracy statistics suggest that direct random-to-isarithm interpolation displays the characteristics of both grid-to-isarithm and random-to-grid interpolation. The various forms of accuracy may also be directly associated with characteristics of the data surfaces under investigation.

a. Geometric accuracy (as expressed by position and elevation errors and isarithm correlation coefficients) is strongly correlated with surface autocorrelation as defined in 'VECT'. Within model type, it is also strongly correlated with NNS.

b. Morphological trueness (as expressed by direction error) is correlated with surface autocorrelation, although the results are strongly associated with those obtained from the use of gridded data.

c. The results closely mimic those obtained from the use

of comparably sized gridded data (between medium and coarse grids).

In general, use of each data type generates its own unique signal in terms of all geometric errors, with the exclusion of Koppe formulae. Use of CONTOUR data produces generally high-accuracy results (see Lindig errors), although it often includes ambiguities (see overlays) and generally interpolates high and produces poor error statistics with generally straight line isarithms. Though use of both SCATTER and BREAKLINE data produces results of a similar accuracy and interpolating slightly low, use of SCATTER data however produces a superior graphical product with fewer ambiguities. The advent of ambiguities from the use of CONTOUR and BREAKLINE data underlines a major problem with such linearly derived data. Use of RIVER data leads to values which are systematically interpolated low; this is an inherent aspect of the data as discussed earlier.

Finally, run costs are summarised in Tables 7.2 and 7.3. While total costs may be quite large as compared with grid-to-isarithm interpolation programs, it should be remembered that half of the total cost is derived from compiling the program and therefore would not exist in a production environment as the compiled form would be stored. Additionally, this cost should, of course, be considered more correctly with reference to the accumulated random-to-grid and grid-to-isarithm costs.

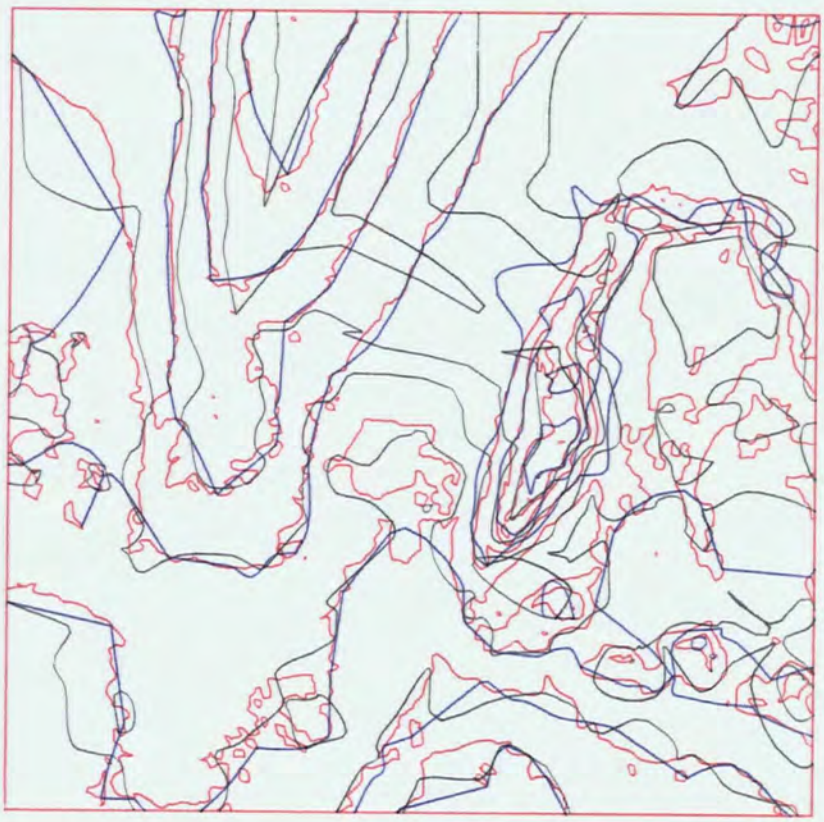
7.4 CONCLUSIONS

Throughout this thesis the two basic methods of isarithm generation from point data (random-to-isarithm interpolation and random-to-grid followed by grid-to-isarithm interpolation) have been kept separate. However, by using Lindig-derived statistics and by considering the relevant overlay plots, these two differing solutions can be considered. For the purpose of this examination, grid-to-isarithm interpolation was performed on the grids interpolated from the point data sets by multiquadric analysis (the overall optimum method of random-to-grid interpolation). In order to provide a similarly smoothed random-to-isarithm plot for comparison, the post-threading smoothing routine of Akima (1974a, 1974b) was utilised (as discussed in A1.4.6) within the random-to-isarithm program DTC. Summary overlay plots (derived from use of the FORV data sets) are shown in Figures 7.27 and 7.28, Koppe plots for each method are shown in Figures 7.29 and 7.30 and the Lindig errors are summarised in Figures 7.31 and 7.32.

From an examination of the overlay isarithm plots with those from the point data (without smoothing), it is clear that the multiquadric (grid-derived) isarithms are more aesthetic and appear of superior accuracy. Small ambiguities in non-smoothed DTC isarithms (as discussed earlier) have become exaggerated when post-threading smoothing is performed, providing several extreme crossing isarithms.

Examining the Koppe plots (Figures 7.29 and 7.30), it is apparent that both methods of isarithm interpolation are

Figure 7.27



FORV : POINT (SMOOTHED)

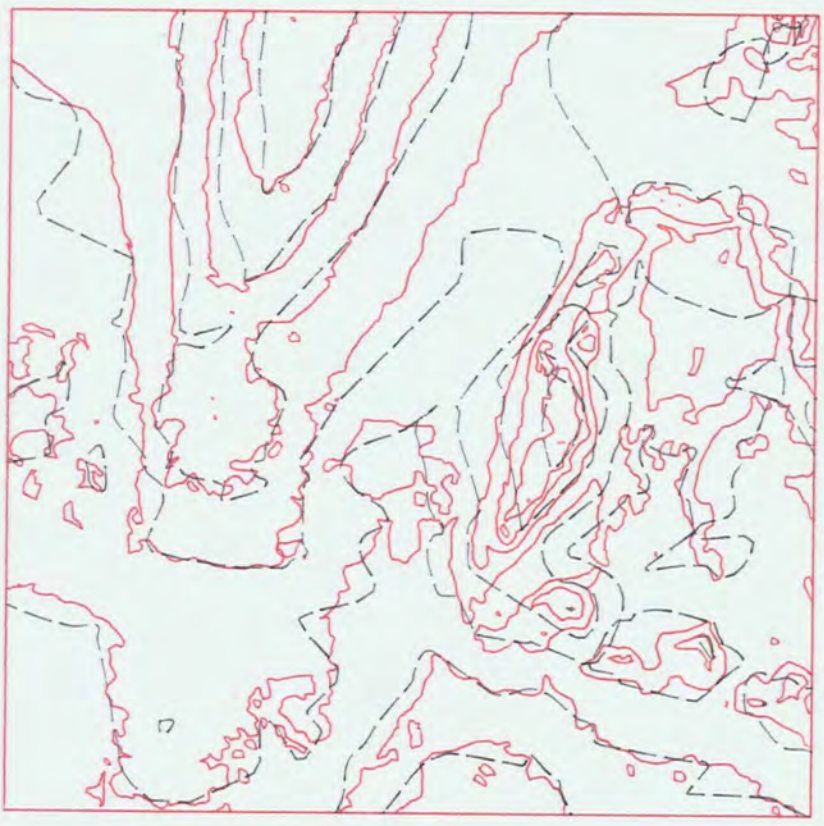
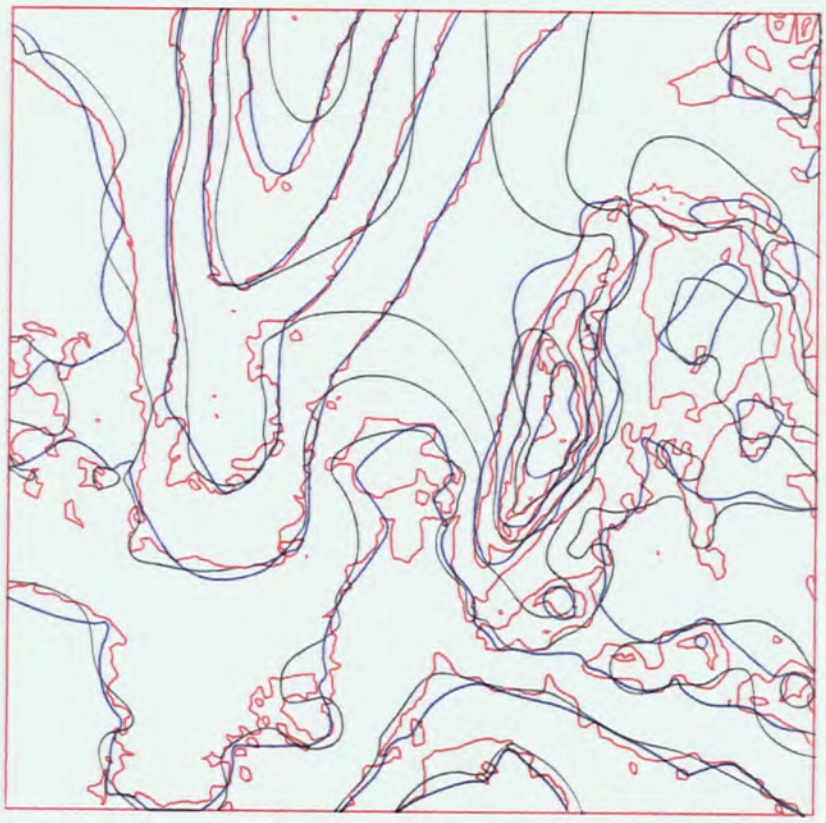
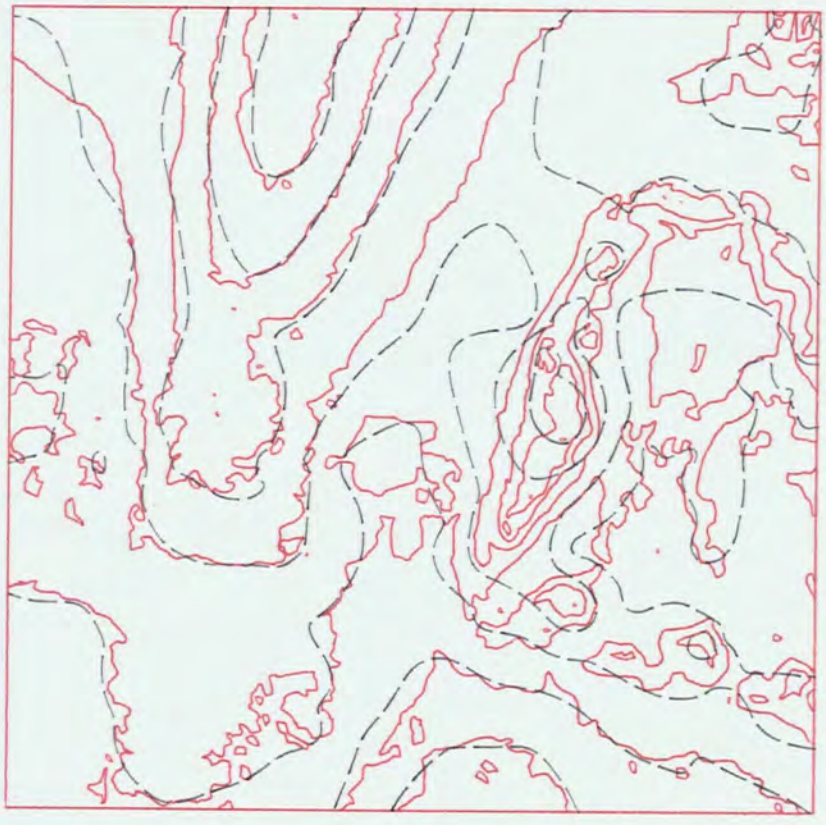
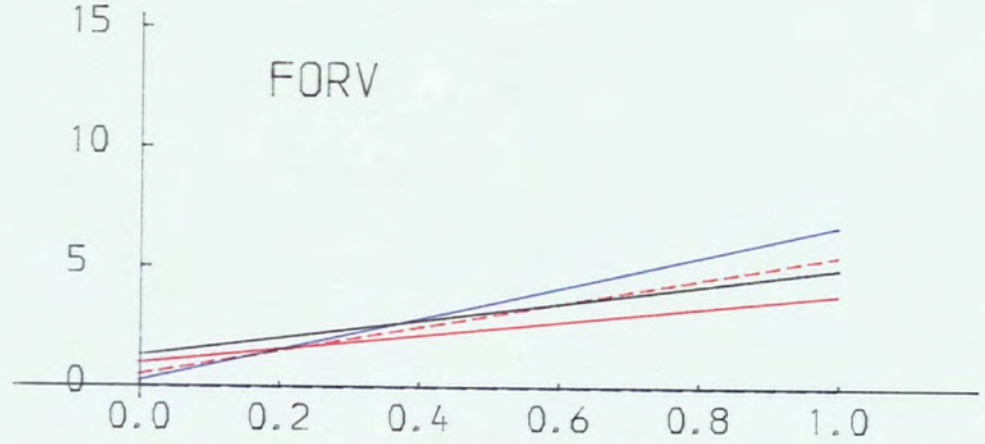


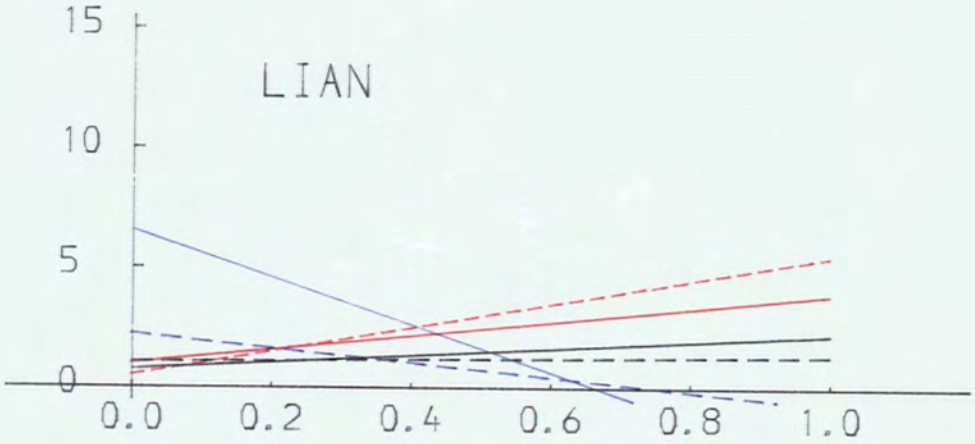
Figure 7.28



FORV : POINT (MULTI)







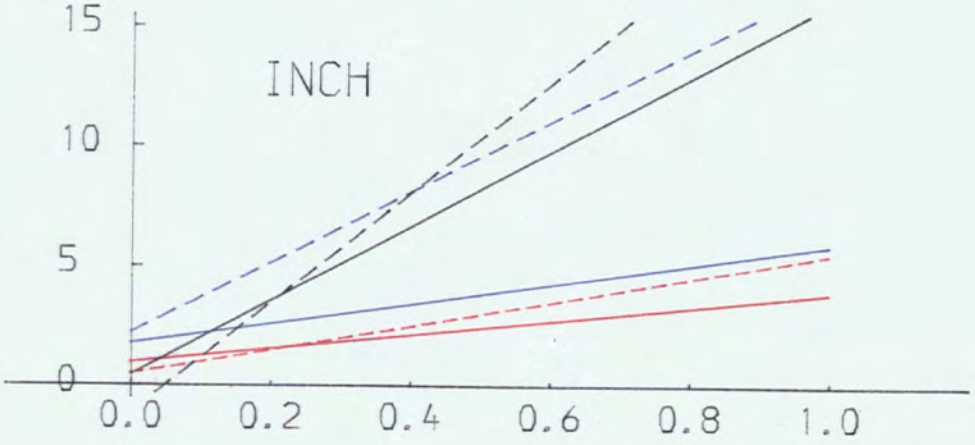
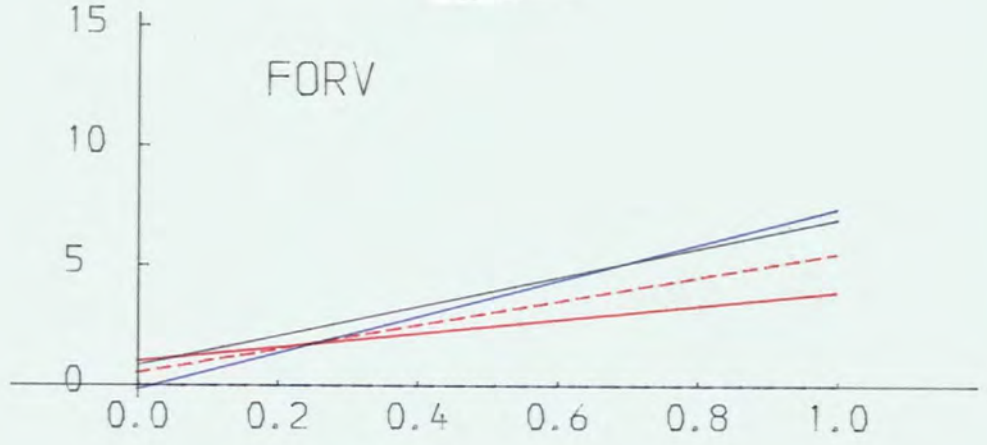
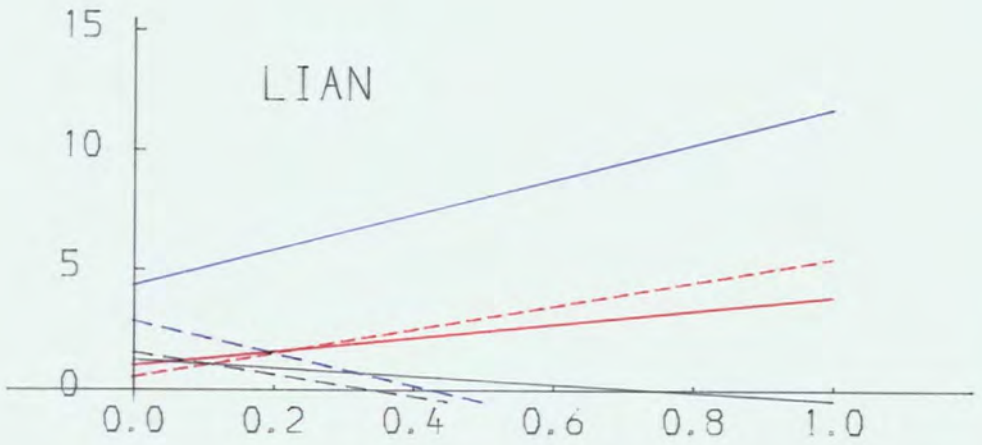


Figure 7.29 Random-to-isarithm (smoothed) Koppe graphs





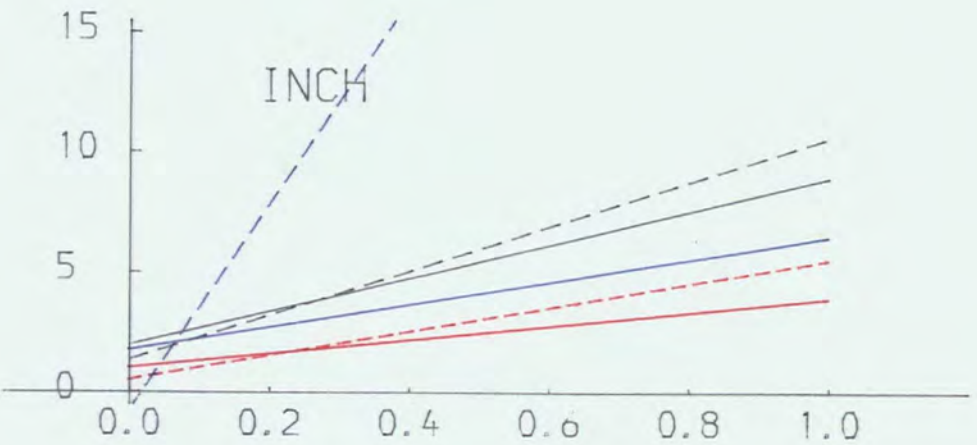


Figure 7.30 Random-to-isarithm (multiquadric derived grid) Koppe graphs

Figure 7.31 Random-to-isarithm (smoothed) Lindig errors

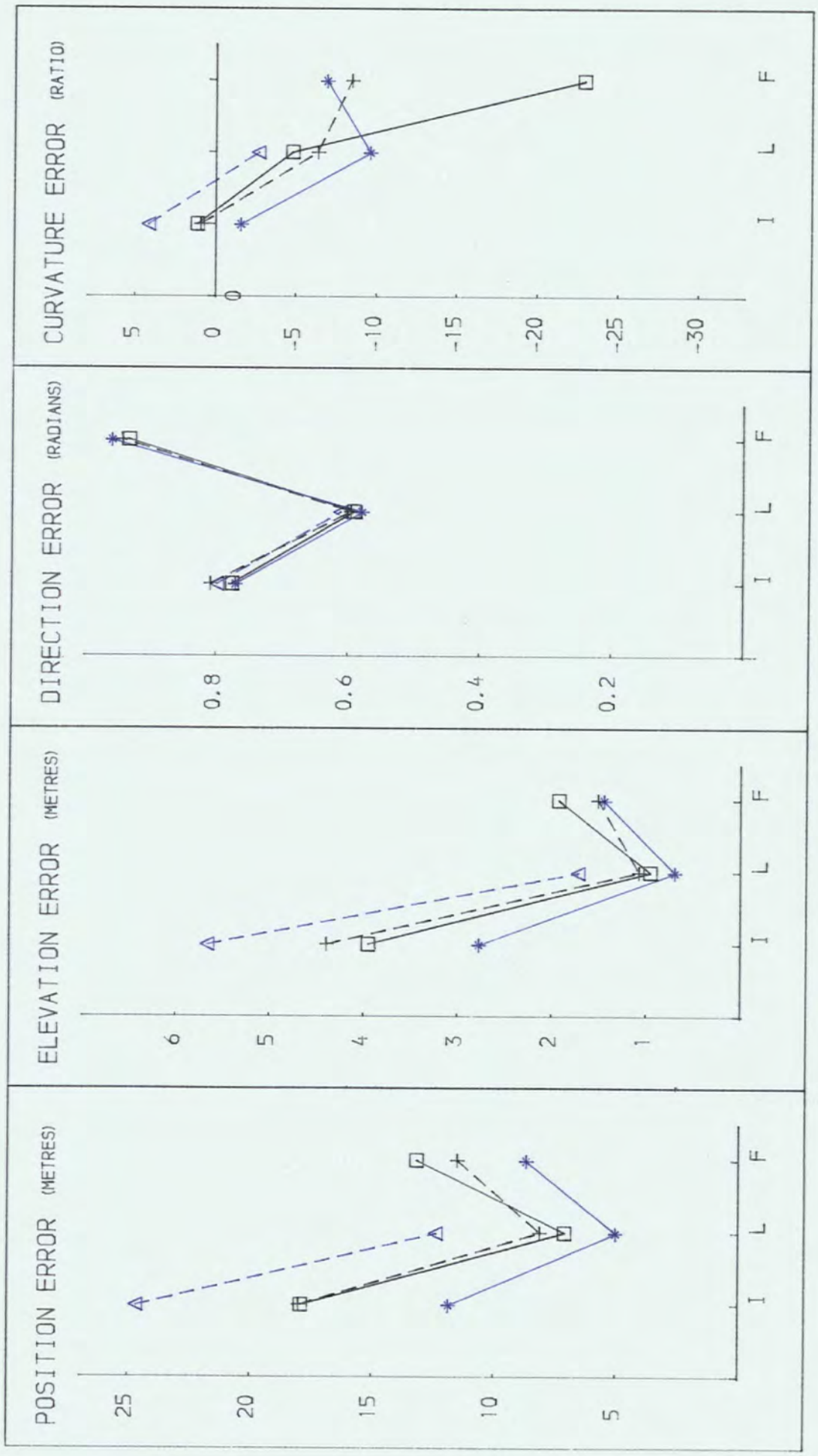
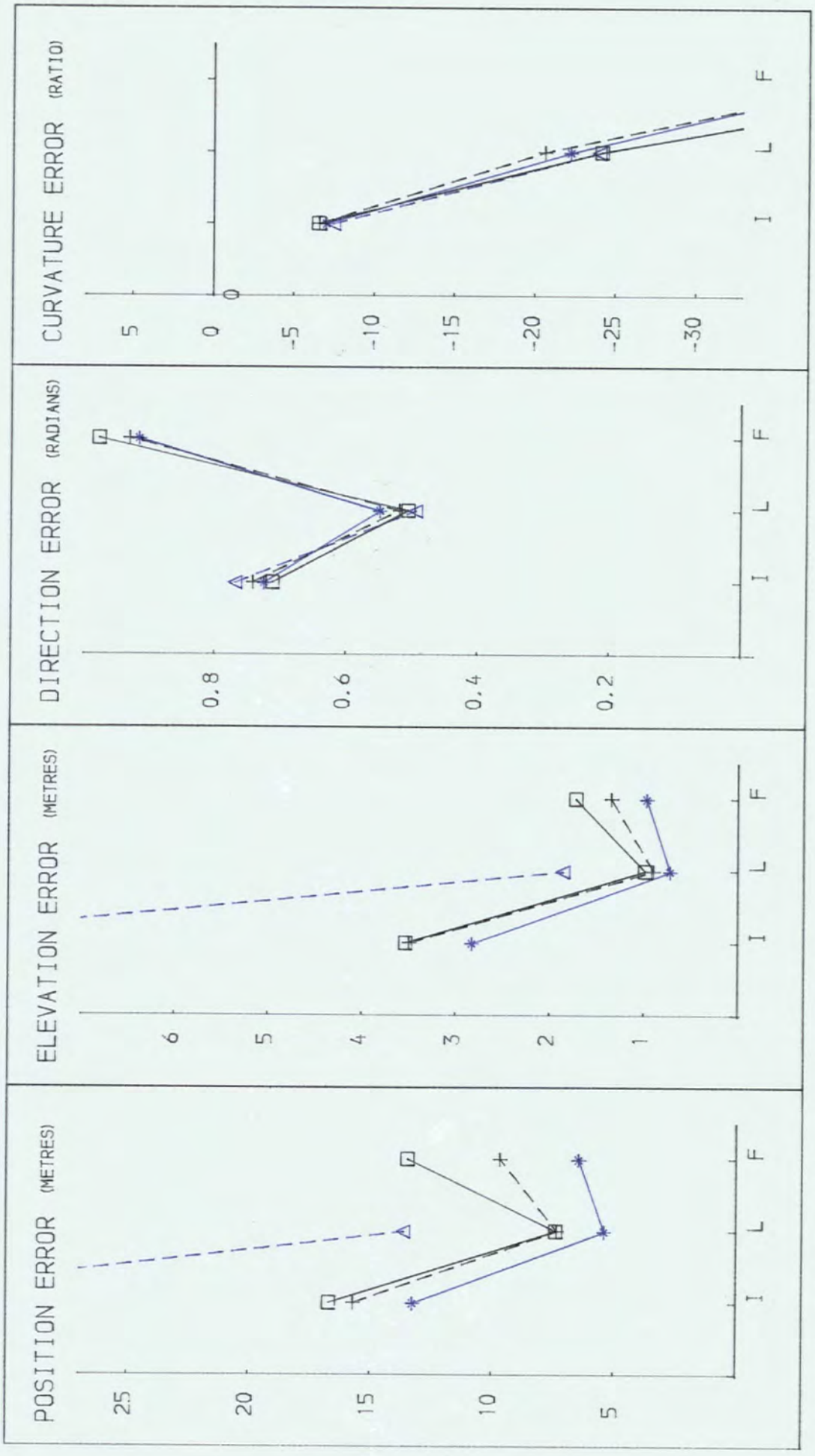


Figure 7.32 Random-to-isarithm (multiquadric derived grid) Lindig errors



consistently inferior to the unsmoothed DTC isarithms. Overall, smoothed random-to-isarithm interpolation performs better with INCH data (macro-features), while the multiquadric-derived isarithms are superior in FORV (micro-features) and to a lesser extent in LIAN.

As regards Lindig errors (Figures 7.31 and 7.32), all three methods generate similar position and elevation errors. Significantly, as established with the grid-to-isarithm interpolation, post-threading smoothing generates isarithms with poorer morphological trueness than without smoothing. Multiquadric-derived isarithms perform comparably to the smoothed isarithms, although the curvature error is significantly poorer.

Therefore, remembering that multiquadric analysis was the optimum method of random-to-grid interpolation, it must be assumed that any other method of random-to-grid interpolation will produce inferior isarithms. While grid-derived isarithms may produce an aesthetically superior product, isarithms generated by random-to-isarithm methods, whether with or without smoothing, have superior accuracy.

CHAPTER 8

CONCLUSION

8 CONCLUSION

8.1 RESULTS

Several conclusions may be drawn from this research with relation to establishing data and surface characteristics, random-to-grid interpolation and isarithm interpolation. While the main purpose of this research was to consider these purely in terms of computer isarithm generation, it has been shown that these aspects are of relevance to multi-disciplinary digital modelling applications (see 2.2).

8.1.1 Data And Surface Characteristics

In evaluating the descriptors, four pertinent points were suggested in 4.4.

a. Surface discrimination. While a summary of all descriptors provides unique identifiers for each surface, these are often conflicting. Autocorrelation descriptors are largely scale-independent and therefore suggest an order of roughness - the surfaces represented by the FORV data (being distinct from those represented by INCH and by LIAN data). The scale-dependent identifiers (surface area and vector dispersion analysis) suggest INCH surfaces to be distinctly rougher than those of LIAN and FORV. Scale-variance suggests a similar order, with those of LIAN and FORV distinct.

b. Data set discrimination. Consistency of data set descriptors decreases with scale dependency. NNS is best, although areal- and to a lesser extent, two-dimensional autocorrelation provide good descriminators. Surface area and vector dispersion analysis are poor and highly variable descriptors.

c. Consistency of the observed and interpolated grid descriptors. In general, there is good consistency between the descriptors derived from the interpolated and similarly-sized observed grids. This allows easy identification of surface type and is noticeably successful with scale-variance analysis.

d. Consistency of the descriptors. This is considerable, with the exception of surface area analysis in which the two sizes of data sets and the gridded data give significantly differing results.

Overall, autocorrelation, scale-variance, NNS and vector dispersion analysis are good descriptors which might also be used to successfully describe non-topographic surfaces and various data sets.

8.1.2 Random-to-grid Interpolation

Various conclusions have been made with regard to

random-to-grid interpolation. It should be remembered that, while these conclusions are considered separately, for the purposes of selection of method or data, they should be considered together.

8.1.2.1 Performance

There are several conclusions to be made with regard to the performance of the various aspects of random-to-grid interpolation.

a. Within pointwise interpolation, there is little difference between the relative performances of various search algorithms. Quadrant-based searches are generally superior, especially for highly correlated data, although a variable radius method is superior for low correlated data. Octant searching is similar to quadrant although, because it finds more distant points, it interpolates poorer results with the data sets considered. Similarly, a simple 'n' nearest neighbour search is poorer than a variable radius search. Whichever search method is used, it must be guaranteed to find a minimum number of points sufficient for the function and fit or else serious problems may result (eg. as in GHOT).

b. In general, weighting functions based solely on distance provide similar results. Where NNSs are similar, autocorrelation functions are useful indicators of the form of such functions although, in general, higher NNSs suggest

the requirement for slower decay weighting functions.

A more significant aspect of weighting is direction weighting, as used in SYMAP. Notwithstanding the poor results from SYMAP (which are probably due to other factors described earlier), its inclusion controls and minimises the effect of poor point distributions, especially clusters and voids.

c. Within pointwise interpolation, where a two-stage interpolation is used involving a weighted average of data surfaces (eg. GPCP), extremes are likely to occur where the NNS is less than about 1.15 and there is low autocorrelation in the data. This is especially so if lineated data are used and may be controlled by constraining the values derived from the data surfaces (eg. as in SYMAP).

d. Generally, quadratic functions are ideal in surface interpolation, eg. multiquadric analysis. However, when used in conjunction with pointwise interpolation, only data sets with a NNS greater than about 1.050 (4-term parabola) and 1.350 (6-term parabola) should be used, and strongly lineated data should be avoided.

e. Kriging is generally a superior method of interpolation where the data distribution is adequate. However, it is a very sensitive method, susceptible to error in areas of extensive data voids.

f. Similarly, patchwise polynomial interpolation is generally an efficient and good interpolator except in areas where data are poorly distributed (e.g. the edge of area under interest and in data voids). Its performance is strongly related to surface morphology.

g. SACM and multiquadric analysis were the most accurate and reliable methods used, mainly because they used simple functions (linear and quadratic respectively) and constrained the functions, throughout, from the data points directly (multiquadric) or by building on well-defined interpolated values (SACM).

h. Global high-order polynomial functions, whilst of use in trend surface analysis, are poor interpolators.

8.1.2.2 Processing Considerations

Associated with the statistical performance of the various interpolation methods are the restrictions caused by any

processing requirements. On considering relative cost and memory requirements, it is apparent that kriging is expensive, while SURFACE II GRAPHICS, patchwise interpolation (either BRAILE or SACM) and even SYMAP are suitable with most well-distributed data sets. Multiquadric analysis, as a result of extensive CPU and memory requirements, should be restricted to work requiring high accuracy results from poorly distributed data.

8.1.2.3 Data Considerations

In addition to the statistical and processing limitations of a particular interpolation method, the effects of the data within the interpolation process must be considered. These can be broadly separated into data distribution properties and surface characteristic properties.

An important data distribution property is described by the NNS. Gridded data (NNS = 2.0) always give superior results. In the data sets used, contour-derived data (NNS = 1.3) and generally evenly scattered data (NNS = 1.15) produce satisfactory results in most cases. Equally, breakline-derived data (NNS = 1.1) and river data (NNS = 0.8) used gave particularly poor results with most interpolation methods, although breakline-derived data can produce satisfactory results with SACM and multiquadric analysis. In addition, the performances of the quadratic and more complex pointwise techniques and of the global and one-stage patchwise techniques are also related to the surface-random/specific nature of the data. Poorly distributed lineated surface-specific data should

be avoided within the procedures tested, especially where directional weighting or searching is not being employed.

In comparison, the surface characteristic properties are of secondary importance. In general, the best optimum interpolation predictors are mostly associated with the magnitude and distribution of local surface gradient (surface area analysis, vector dispersion analysis, pointwise polynomial - gradient). However, for some of the more specialised search methods (OCTANT) and function types (quadratic and kriging), the autocorrelation functions provide excellent predictors.

8.1.3 Grid-to-isarithm Interpolation

Various conclusions have been realised with regard to grid-to-isarithm interpolation.

8.1.3.1 Accuracy Prediction

Having investigated the various methods of expressing the accuracy of isarithms, it is apparent that the methods of predicting the accuracy of isarithms vary considerably.

a. The two-dimensional autocorrelation function is an excellent predictor of the accuracy as expressed by scale-independent parameters (i.e. morphological trueness - direction error - and geometric accuracy - isarithm correlation).

b. Vector dispersion analysis is a good predictor of morphological trueness - curvature error.

c. Surface area analysis is a good predictor for geometric accuracy - position and elevation error.

d. Koppe formulae are highly variable and therefore difficult to use to predict accuracy.

Therefore, irrespective of the form of the accuracy statement, a good prediction of the relative performances of the data may be established in advance by considering the characteristic of the data sample.

8.1.3.2 Optimum Method

Three considerations must be discussed when deciding which is the optimum method of grid-to-isarithm interpolation.

a. In terms of accuracy, isarithm interpolation utilising patchwise polynomial functions (eg. GPCP) is consistently superior (with the possible exclusion of curvature error). Excluding this, centre-point triangulation algorithms (eg. CONSYS and GINOSURF) are superior with smooth surfaces and diagonal triangulation (eg. GHOST) and direct grid schemes (eg. SACM and SURFACE II GRAPHICS) are superior with rough surfaces.

Post-threading isarithm smoothing does not necessarily produce more- or less-accurate isarithms (eg. CONSYS against GINOSURF and GHOST against SURFACE II GRAPHICS).

b. As regards cost, all methods require similar expenditure with the exclusion of GPCP - requiring approximately two to three times the cost of the other packages, although this is partly because of associated costs (data reformatting). An additional consideration is that of post-threading smoothing - which increases the cost by approximately 50%.

c. Finally, the logistics of using each method must be considered. SACM and SURFACE II GRAPHICS are easiest to use since only a short command file must be established - data being input in a user-specified format. CONSYS, GHOST and GINOSURF are subroutine libraries and therefore require a calling program. GPCP is the most difficult package to use, since it requires precise data reformatting in conjunction with a command file.

8.1.4 Random-to-isarithm Interpolation

Random-to-isarithm interpolation produces the equivalent of both random-to-grid and grid-to-isarithm interpolations.

a. Geometric accuracy is strongly correlated with both

surface autocorrelation and the NNS for each of the data sets.

b. Morphological trueness - direction error - is correlated with surface autocorrelation in a fashion very similar to that derived for gridded data.

c. Contour-derived data produce better results than scatter- and breakline-derived data. River-derived data produce poor results.

d. Costs are similar to those required for isarithm interpolation from full-grids in GPCP (there being ten times more points in the grid than in the data point sample). However, the true cost of the actual program run is similar to the other grid-to-isarithm runs, which represent only half the total random-to-isarithm generation process.

e. Direct random-to-isarithm interpolation would appear to produce isarithms at least as good as random-to-grid followed by grid-to-isarithm interpolation, although not of such a high aesthetic quality.

8.2 FUTURE RESEARCH

From the previous discussion, six topics for future research may be suggested.

a. Only three surfaces have been examined within this thesis; more should be involved to provide a firmer basis for prediction.

b. Similarly, only a few data types have been considered. The effects of using highly linear data and especially irregular, approximate and progressive grids should be considered.

c. The present study considers the data sets and surfaces within the context of their discipline, relative dimensions and scale. The effects of removing dimensionality completely may be easily studied within the established framework of research.

d. Although it was originally intended to consider breaklines within interpolation as a morphological boundary, it soon became clear that the author did not have access to appropriate software. The importance of using lines or polygons to break up the surface into morphologically significant areas prior to interpolation should therefore be examined.

e. As regards interpolation methods examined, a major technique not fully examined was that of splining, especially within patchwise random-to-grid interpolation.

f. The importance of patchwise interpolation, multiquadric analysis, direction weighting and restraining extremes has been shown within the random-to-grid interpolation. These aspects should be combined within one efficient interpolation package. It should be feasible to make this package choose defaults on the basis of the prior contemplation of the data characteristics.

BIBLIOGRAPHY

BIBLIOGRAPHY

Ackermann, D.L. and Crombie, M.A. (1975) Computing a line-of-sight using digital image matching and analytical photogrammetry. US Army Engineering Topographic Laboratories, Fort Belvoir, Virginia.

Ackermann, F. (1978) Experimental Investigation into the accuracy of contouring from DTM. Photogrammetric Engineering and Remote Sensing 44, 12, 1537-48.

Adams, J.M. (1969) Mapping with a third dimension. Geographical Magazine 42, 1, 45-9.

Agterberg, F.P. (1969) Interpolation of areally distributed data. Quarterly of the Colorado School of Mines 64, 3, 217-37.

Ahlberg, J.H. , Nilson, E.N. and Walsh, J.L. (1967) The theory of splines and their applications. Academic Press, New York.

Akima, H. (1970) A new method of interpolation and smooth curve fitting based on local procedures. Journal of the ACM 17, 4, 589-602.

Akima, H. (1974a) A method of bivariate interpolation and smooth surface fitting based on local procedures. Communications of the ACM 17, 1, 18-20.

Akima, H. (1974b) Bivariate interpolation and smooth surface fitting based on local procedures. Communications of the ACM 17, 1, 26-31.

Allam, M.M. (1978) DTM's application in topographic mapping. Proceedings of the Digital Terrain Models (DTM) Symposium St. Louis, Missouri, 1-15.

Alster C.J. (1960) Checking vertical accuracy of topographic maps. Surveying and Mapping 20, 4, 446-8.

American Society of Photogrammetry (1965) Manual of Photogrammetry. (3rd edition), Banta, Mensha.

Anda, B. (1976) Automatic hill shading with a photohead. ITC Journal 3, 546-53.

Applied Research Cambridge (1979a) Systems for planning. brochure.

Applied Research Cambridge (1979b) CIIS- Contouring interpolation and integration of surfaces. brochure.

Arthur, D.W.G. (1965) Interpolation of a function of many variables. Photogrammetric Engineering 31, 2, 348-9.

Arthur, D.W.G. (1973) Interpolation of a function of many variables, II. Photogrammetric Engineering 39, 3, 261-6.

Assiter, E. (1967) Surface approximations and contour mapping - user's guide. Applications Consultants Inc., Houston, Texas.

Assmus, E. (1974) Extension of the Stuttgart contour program to treating terrain break-lines. Proceedings of the Symposium of Commission III, I.S.P. Stuttgart.

Assmus, E. (1976) Extension of Stuttgart contour program to treating terrain break-lines - theory and results. XIIIth Congress of the International Society for Photogrammetry Helsinki.

Assmus, E. and Stanger, W. (1978) A general digital terrain model and some applications in road construction. Research paper.

Ayeni, O. (1976) Optimum sampling for digital terrain models. XIIIth Congress of the International Society for Photogrammetry Helsinki.

Ayeni, O.O. (1979) Optimum least squares interpolation for DTMs. Photogrammetric Record 9, 53, 633-44.

Barrett A.N. and Rhind, D.W. (1974) Retrospective application of boundary constraints in automatic contouring by digital spatial filtering. Journal of the International Association for Mathematical Geology 6, 1, 83-9.

Batcha, J.P. and Reese, J.R. (1965) Surface determination and automatic contouring for mineral exploration extraction, and processing. Quarterly of the Colorado School of Mines 49, 4, 1-14.

Baxter, R.S. (1976) Computer and statistical techniques for planners. Methuen, London.

Beckman, P. and Spizzichino, A. (1963) The scattering of electromagnetic waves from rough surfaces. International Series of Monographs on Electromagnetic Waves (4), Pergamon, London.

Bell, J.F. (1977) The development of the Ordnance Survey 1:25,000 scale derived digital map. The British Cartographic Society Annual Symposium Durham.

Berztiss, A.T. (1964) Least squares fitting of polynomials to irregularly spaced data. SIAM Review 6, 3, 203-27.

Bethel, J., Crawley, B.G., Shepphird G. and Hussain, M. (1978) The automatic generation and utilization of terrain surface data on a production basis. Proceedings, American Society of Photogrammetry 4th Annual Meeting, Washington, 301-19.

Bhattacharyya, P.J. and Ross, D.I. (1972) Mid-Atlantic ridge near 45°N: Computer interpolation and contouring of the bathymetry and magnetics. Marine Science Paper 11, Information Canada, Ottawa.

Bickmore, D.P. (1967) The Oxford Cartographic Data Bank. Proceedings, Commonwealth Survey Officers Conference Cambridge, Paper J9.

Blachut, T.J. and van Wijk, M.C. (1976) Results of the international orthophoto experiment, 1972-76. Photogrammetric Engineering and Remote Sensing 42, 12, 1483-98.

Blakney, W.G.G. (1968) Accuracy standards for topographic mapping. Photogrammetric Engineering and Remote Sensing 34, 10, 1040-2.

Blaschke, I.W. (1968) Digital terrain model - A presentation of the practical application of some existing DTMs. XIth Congress of the International Society for Photogrammetry Lausanne.

Blaschke, W. (1969) Problems in obtaining and processing of photogrammetric data for digital terrain models. XIth Congress of the International Society for Photogrammetry Lausanne.

Bolondi, G., Rocca, F. and Zanoleti, S. (1976) Automatic contouring of faulted subsurfaces. Geophysics 41, 6, 1377-93.

Botman, A.G. and Kubik, K.K.Th. (1979) On the theoretical accuracy of the moving average method for surface estimation. ITC Journal 1, 68-85.

Boyle, A.R. (1971) Automatic cartography: Special problems of hydrographic charting. International Hydrographic Review 48, 2, 85-92.

Braile, L.W. (1978) Comparison of four random to grid methods. Computers and Geosciences 4, 4, 341-9.

Braker, F. (1975) Generation of irregular point grids from contour lines for digital terrain models. MSc.(Ph.E.) thesis, ITC, Enschede.

Brassel, K. (1974) A model for automatic hill shading. The American Cartographer 1, 1, 15-27.

Briggs, I.C. (1974) Machine contouring using minimum curvature. Geophysics 39, 1, 39-48.

Brooks, P.D. (1970) An investigation of the accuracy of multiquadric equations of topography. MSc. thesis, Iowa State University.

Calcomp (1973) GPCP - A general purpose contouring program : User's manual. California Computer Products, Inc., Anaheim, California.

Caswell, F. (1982) Success in statistics. John Murray, London.

Cederquist, G.N. (1976) CONSYS : A collection of Fortran subroutines to produce contour maps of data surfaces defined on regular grids. Cooley Electronics Lab. Michigan University, Technical memo. C.E.L.-TM-112,013514-2-M.

Chiles, J.P. and Chauvet, P. (1975) Kriging : A method for cartography of the sea floor. International Hydrographic Review 52, 25-41.

Chorley, R.J. and Haggett, P. (1965) Trend-surface mapping in geographical research. Transactions, Institute of British Geographers 37, 47-67.

Clerici, E. and Kubik, K. (1974) The theoretical accuracy of point interpolation on topographic surfaces. Proceedings of the Symposium of Commission III I.S.P. Stuttgart.

Cliff, A.D. (1970) Computing the spatial correspondence between geographical patterns. Transactions, Institute of British Geographers 50, 143-54.

Cliff, A.D. and Ord, J.K. (1973) Spatial autocorrelation. Pion, London.

Cline, A.K. (1974a) Scalar- and planar- valued curve fitting using splines under tension. Communications of the ACM 17, 4, 218-20.

Cline, A K. (1974b) Six subprograms for curve fitting using splines under tension. Communications of the ACM 17, 4, 220-3.

Cole, A.J. (1969) An iterative approach to the fitting of trend surfaces. Kansas Geological Survey Computer Contribution, no.37 Lawrence, Kansas.

Cole, A.J. , Jordan, C. and Merriam, D.F. (1967) Fortran II program for progressive linear fit of surfaces on a quadratic base using an IBM 1620 computer. Kansas Geological Survey Computer Contribution, no.15, Lawrence, Kansas.

Collins, S.H. and Moon G.C. (1981) Algorithms for dense digital terrain models. Photogrammetric Engineering and Remote Sensing 47, 1, 71-6.

Connelly, D.S. (1971) An experiment in contour map smoothing on the ECU automated contouring system. The Cartographic Journal 8, 1, 59-66.

Conte, S.D. and de Boor, C. (1972) Elementary numerical analysis. McGraw-Hill Kogakusha Ltd. Tokyo.

Cook, L.T., Dwyer, S.J., Batnitsky, S. and Lee, K.R. (1983) A three-dimensional display system for diagnostic imaging applications. IEEE Computer Graphics and Applications 3, August, 13-9.

Court, A. (1970) Map comparisons. Economic Geography 46, 2, 435-8.

Cox, N. (1979) Models and methods in hillslope profile morphometry. PhD. thesis, University of Durham.

Crain, I.K. (1970) Computer interpolation and contouring of two-dimensional data A review. Geoexploration 8, 71-86.

Crain, I.K. and Bhattacharyya, B.K. (1967) Treatment of non-equispaced two-dimensional data with a digital computer. Geoexploration 5, 173-94.

Craine, G.S. and Heatherington, S. (1977) Ground model accuracy. Paper presented at the PRTC Summer Annual Meeting.

Craine, G.S., Houlton, J.M. and Malcomson, E. (1974) MOSS, A modelling system for highway design and related disciplines. Public Works Congress 317-41.

Cubitt, J.M. and Celenk, O. (1976) Fortran program for producing stereograms in geology. Computers and Geosciences 1, 3, 207-11.

CAD Centre (1976) Gino-F. Computer Aided Design Centre, Cambridge.

Davies, R.S. (1980) A review of computer techniques for representation of geometry. Proceedings of the Computer Aided Design Conference Brighton 213-25.

Davis, J.C. (1973) Statistics and data analysis in geology. John Wiley and Sons, Inc. New York.

Davis, J.C. (1975) Contouring algorithms. Proceedings of the International Symposium on computer-assisted cartography". AUTO-CARTO II, ACSM Cartography division.

Davis, M.W.D. and David, M. (1978) Automatic kriging and contouring in the presence of trends. The Journal of Canadian Petroleum Technology 17, 1, 90-99.

Dayhoff, M.O. (1963) A contour map program for X-ray crystallography. Communications of the ACM 6, 10, 620-2.

Doyle, F.J. (1978) Digital terrain models: an overview. Photogrammetric Engineering and Remote Sensing 44, 12, 1481-5.

Drewry, D.J. (1975) Terrain units in Eastern Antarctica. Nature 256, 194-5.

Dutton, G. (1982) Land alive: An algorithm for four-color hill-shading enables realistic and colorful relief-maps to be produced at a display terminal. Perspectives in Computing 2:26-39.

Ebner, H., Hofmann-Wellenhof, B., Reiss, P. and Steidler, F. (1980) HIFI - A minicomputer package for height interpolation by finite elements. Congress of the International Society for Photogrammetry Hamburg.

Edwards, K.A. (1972) Estimating areal rainfall by fitting surfaces to irregularly spaced data. Proceedings, World Meteorological Office/IAHS Symposium 2, 565-87.

Elfick, M.H. (197) Contouring by use of a triangular mesh. Cartographic Journal 16, 1, 24-9.

Evans, I.S. (1972) General geomorphometry, derivatives of altitude, and descriptive statistics In Chorley, R.J. (Ed), Spatial analysis in geomorphology. Methuen, London, 17-90.

Evans, I.S. (1979) An integrated system of terrain analysis and slope mapping. Dept. of Geography, Univ. of Durham, Grant DA-ERO-591-73-G0040, final report.

Evans, I.S. and Bain, I. (1974a) Statistical characterisation of altitude matrices by computer. Dept. of Geography, Univ. of Durham, Grant DA-ERO-591-73-G0040, first progress report.

Evans, I.S. and Bain, I. (1974b) Statistical characterization of altitude matrices by computer. Dept. of Geography, Univ. of Durham, Grant DA-ERO-591-73-G0040, second progress report.

Falconer, K.J. (1971) A general purpose algorithm for contouring over scattered data points. National Physical Laboratory, Report NAC6.

Feder, J, , and Freeman, H. (1966) Digital curve matching using a contour correlation algorithm. IEEE International Convention Record part 3, 69-85.

Fisher, R.A. (1953) Dispersion on a sphere. Proceedings Royal Society Series A, 217, 295-305.

Forsythe, G.E. (1957) Generation and use of orthogonal polynomials for data fitting with a digital computer, Journal of the Society for Industrial Applications of Maths 5, 2, 74-88.

Franke, R. (1979) A critical comparison of some methods for interpolation of scattered data. Naval Postgraduate School NPS-53-79-003.

Frye, W.H. (1978) A plan for comparing the triangulated networks (TIN) and uniform grid (URG) methods of digital terrain modelling. Naval Warfare research center research memo. Cont. No.0014-77-C-0698.

Fullenwider, D.R., Lefever, J.P. and Martin, A.C. (1981) Computer graphics and the practice of architecture. IEEE Computer Graphics and Applications 1, October, 19-26.

Gimbert, A.W. and Cubitt, R.E. (1973) Automatic map drawing using a computer line printer. GLC Intelligence Unit Quarterly Bulletin 25, 53-61.

Gold, C.M., Charlesworth, H.A.K. and Kilby, W.E. (1981) Coal resource evaluation in deformed sequences, using digital terrain models. Bulletin of Canadian Petroleum Geology, 29, 2, 259-66.

Gordon, N.D. (1981) Computer graphics in the New Zealand Meteorological Service. New Zealand Cartographic Journal 11, 1, 6-14.

Grassie, D.N.D. (1982a) Contouring by computer: some observations. In British Cartographic Society, Computers in cartography. Special Publication 2, 93-116.

Grassie, D.N.D. (1982b) Random-to-grid-interpolation and contouring in digital terrain modelling. Royal Society Digital Mapping Sub-committee Paper presented to a meeting on digital terrain models.

Green, P.J. and Sibson, R. (1978) Computing Dirichlet tesselations in the plane. The Computer Journal 21, 2, 168-73.

Greville, T.N.E. (1967) Spline functions interpolation, and numerical quadrature. In Ralston, A, and Wilf, H.S. Mathematical methods for digital computers. Wiley, New York, 2, 156-68.

Greville, T.N.E. (1969) Introduction to spline functions. In Greville, T.N.E. Theory and applications of spline functions. Academic Press, New York, 1-35.

Grist, M.W. (1972) Digital ground models : An account of recent research. Photogrammetric Record 7, 40, 424-41.

Guralnik, D B. (1978) Concise English dictionary. Collins, Glasgow.

Gustafson, G.C. (1977) Koppe errors on automatic contouring. Technical papers of the ACSM 37th meeting 77-232, 305-17.

Hall, J.K. (1976) Algorithms and programs for the rapid computation of area and center of mass. Computers and Geosciences 1, 3, 203-5.

Hammond, E.H. (1958) Procedures in the descriptive analysis of terrain. Project no. NR 387-015 Contract No. Nonr-1202(01).

Hardy, R.L. (1971) Multiquadric equations of topography and other irregular surfaces. Journal of Geophysical Research 76, 8, 1905-15.

Hardy, R.L. (1972a) Analytical topographic surfaces by spatial intersection. Photogrammetric Engineering 38, 5, 452-8.

Hardy, R.L. (1972b) Geodetic applications of multiquadric analysis. Algemeine Vermessungs Nachrichten 79, 10, 398-406.

Hardy, R.L. (1975) Research results in the application of multiquadric equations to surveying and mapping problems. Surveying and Mapping 35, 4, 321-32.

Harley, J.B. (197) Ordnance Survey maps A descriptive manual. The Ordnance Survey, Southampton.

Harrington, G.J. (1976) Contour plans drawn by computer using random stadia survey data. New Zealand Surveyor 28, 495-504.

Harrison, J.C. (1978) The LAMIS system. RICS Land Surveyors Division and North East London Polytechnic, "Databanks and Digital Mapping".

Hartley, P.J. and Judd, C.J. (1980) Curve and surface representations for Bezier B-spline systems. Proceedings of the Computer Aided Design Conference Brighton, 226-36.

Hayes, J.G. and Halliday, J. (1972) The least-squares fitting of cubic spline surfaces to general data sets. NPL report NAC22.

Heap, B.R. (1972) Algorithms for the production of contour maps over an irregular triangular mesh. National Physical Laboratory, Report NAC10.

Heap, B.R. and Pink, M.G. (1969) Three contouring algorithms. National Physical Laboratory, Report DNAM81.

Henley, S. (1976) Autocorrelation coefficients from irregularly spaced data. Computers and Geosciences 2, 4, 437-8.

Hobson, R.D. (1967) Fortran IV programs to determine surface roughness in topography for the CDC 3400 computer. Kansas State Geological Survey Computer Contribution, no.14, Lawrence, Kansas.

Hobson, R.D. (1972) Surface roughness in topography : Quantitative approach. In Chorley, R.J. (Ed.), Spatial analysis in geomorphology. Methuen, London, 221-45.

Holroyd, M.T. and Bhattacharyya, B.K. (1970) Automatic contouring of geophysical data using bicubic spline interpolation. Geological Survey of Canada, Paper no. 70-55.

Howard, S.M. (1968) A cartographic data bank for Ordnance Survey maps. The Cartographic Journal, 5, 1, 48-53.

Howarth, R.J. (1969) A review of trend surface analysis in geology. Proceedings, PTRC Seminar 5-10.

Huijbregts, C. and Matheron, G. (1970) Universal Kriging. Canadian Institute of Mining and Metallurgy Special volume 12, 159-169.

Imhof, E. (1965) Kartographische Gelaendedarstellung. Walter de Gruyter & Co., Berlin.

Institute of Civil Engineers (1977) Surface modelling by computer. Proceedings of a confererence jointly sponsored by the RICS. and ICE. London.

Irwin, M.St.G. (1971) Developments in automated cartography at the Ordnance Survey. The Cartographic Journal 8, 2, 133-8.

James, W.R. (1966) Fortran IV program using double Fourier series for surface fitting of irregularly spaced data. Kansas Geological Survey, Computer Contribution, no.5, Lawrence, Kansas.

Jancaitis, J.R. (1978) Elevation data compaction by polynomial modelling. U.S. Army Engineer Topographic Laboratories, Fort Belvoir, Virginia.

Jancaitis, J.R. and Moore, W.R. (1978) Near real time - Application of digital terrain data in minicomputer environment. US Army Engineer Topographic Laboratories, Fort Belvoir,VA 22060.

Kalkani, E.C. (1977) Evaluation technique to determine relative accuracies of contour maps. Geophysics 42, 4, 860-7.

Kane, V.E. , Begovich, C.L. , Butz, T.R. and Myers, D.E. (1982) Interpretation of regional geochemistry using optimal interpolation parameters. Computers and Geosciences 8, 2, 117-35,

Keates, J.S. (1973) Cartographic design and production. Longman, London.

Keir, K.M. (1969) Photogrammetry and current practice in road design. XIth Congress of the International Society for Photogrammetry Lausanne.

Keir, K.M. (1976) A digital mapping system designed for the commercial air survey market. XIIIth Congress of the International Society for Photogrammetry Helsinki.

Kendall, M.G. and Buckland, W.R. (1982) A dictionary of statistical terms. Longman Group Ltd, London.

Koppe, C. (1902) Die neue topographische Landeskarte des Herzogtums Braunschweig in Massstabe 1:10,000. Zeitschrift fuer Vermessungswesen 31, 14, 397-424.

Kraus, K. (1973) A general digital terrain model - Theory and application. In Ackermann, F. , Numerische photogrammetrie. Sammlung Wichmann, Neue Folge Buchreihe Band 5. Karlsruhe.

Krige, D.G. (1966) A study of gold and uranium distribution patterns in the Klerksdorp gold field. Geoexploration 4, 43-53.

Kuester, J.L. and Mize, J.H. (1973) Optimization techniques with Fortran. McGraw-Hill, New York.

Lam, N.S-N. (1983) Spatial interpolation methods: a review. The American Cartographer 10, 2, 129-49.

Lang, T. (1969) Rules for robot draughtsmen. Geographical Magazine 42, 1, 50-1.

Leberl, F. (1973) Interpolation in square grid DTM. ITC Journal 5, 756-807.

Liebenberg, E. (1976) SYMAP : Its uses and abuses. The Cartographic Journal 13, 1, pp.26-36.

Lindig, G. (1956a) Neue methoden der Schichtlinienprufung. Zeitschrift für Vermessungswesen 7, 244-51.

Lindig, G. (1956b) Neue methoden der Schichtlinienprufung. Zeitschrift für Vermessungswesen 8, 296-303.

Lips, M. (1964) Investigation of Photogrammetric contouring of forested areas for small scale maps. Xth Congress of the International Society for Photogrammetry .

Lynn, P.P. (1975) Rainfall interpolation using multiquadric surfaces. Computer Applications 2, 3-4, 321-34.

Makarovic, B. (1972) The applications of automated equipment in photogrammetry. ITC Publications A53.

Makarovic, B. (1973) Progressive sampling for digital terrain models. ITC Journal 3, 397-416.

Makarovic, B. (1974) Conversion of fidelity into accuracy. ITC Journal 4, 506-17.

Makarovic, B. (1975) Amended strategy for progressive sampling. ITC Journal 1, 117-28.

Makarovic, B. (1976) A digital terrain model system. ITC Journal 1, 57-83.

Makarovic, B. (1979) From progressive to composite sampling for digital terrain models. Geo-Processing 1, 145-66.

Malcomson, E. (1973) MOSS, Digital models by computer. Burisa 5, 16-7.

Marckwardt, W. (1978) The accuracy of orthophotos and simultaneously collected terrain height data. Photogrammetric Engineering and Remote Sensing 44, 5, 575-8.

Marino, J.S (1979) Identification of characteristic points along naturally occurring lines / An empirical study. The Canadian Cartographer 16, 1, 70-80

Mark, D.M. (1975a) Computer analysis of topography : A comparison of terrain storage methods. Geografiska Annaler 57A, 3-4, 179-88.

Mark, D.M. (1975b) Geomorphometric parameters - A review and evaluation. Geografiska Annaler 57A, 3-4, 165-70.

Mark, D.M. (1979) Phenomenon-based data-structuring and digital terrain modelling. Geo-processing 1, 27-36.

Marsden, L.E. (1960) How the National map accuracy standards were developed. Surveying and Mapping 20, 4, 427-39.

Marsik, Z. (1971) Automatic relief shading. Photogrammetria 27, 2, 57-70.

Matheron, G. (1971) The theory of random variables and its applications. Les Cahiers du Centre de Morphologie Mathematique de Fontainebleau, no.5.

Maxwell, D.A. (1970) Mathematical surface approximation of the terrain. Highway Research Record 319, 16-29.

McConalogue, D.J. (1970) A quasi-intrinsic scheme for passing a smooth curve through a discrete set of points. The Computer Journal 13, 4, 392-6.

McCullagh, M.J. (1978) MAPCAT - The development of a mini computer graphics system Computer Applications 5, 1-2, 831-8.

McCullagh, M.J. (1983) Transformation of contour strings to a rectangular grid based digital elevation model. Proceedings, Euro-Carto II .

McCullagh, M.J. and Ross, C.G. (1980) Delaunay triangulation of a random data set for isarithmic mapping. The Cartographic Journal 17, 2, 93-9.

McCullagh, M.J. and Sampson, R.J. (1972) User desires and graphics capability in the academic environment. The Cartographic Journal 9, 2, 109-22.

McDonald, M.F. and Katz, E.J. (1969) Quantitative method for describing the regional topography of the ocean floor. Journal of Geophysical Research 74, 10, 2597-607.

McIntyre, D.B., Pollard, D.D. and Smith, R. (1968) Computer programs for automated contouring Kansas Geological Survey Computer Contribution, no.23 Lawrence, Kansas.

McLain, D.H. (1976) Two-dimensional interpolation from random data. The Computer Journal 19, 2, 178-81.

Menmuir, P. (1974) The automatic producion of weather maps. The Cartographic Journal 11, 1, 12-8.

Merriam, D.F. and Sneath, P.H.A. (1966) Quantitative comparison of contour maps. Journal of Geophysical Research 71, 4, 1105-15.

Miller, C.L. and Laflamme, R.A. (1958) The digital terrain model - Theory and application. Photogrammetric Engineering 24, 3, 433-42.

Moellering, H. and Tobler, W. (1972) Geographical variances. Geographical Analysis 4, 1, 34-50.

Monahan, D. (1971) The limitations of automatically processing bathymetric data. In International Cartography Association, Automation in Cartography. ICA, Paris, 346-52.

Monmonier, M.S. (1971) Digitized map measurement and correlation applied to an example in crop ecology. The Geographical Review 61, 1, 51-71.

Monmonier, M.S., Pfaltz, J.L. and Rosenfeld, A. (1966) Surface area from contour maps. Photogrammetric Engineering 32, 3, 476-82.

Moore, R.F. and Simpson C.J. (1982) Computer manipulation of a digital terrain model (DTM) of Australia. BMR Journal of Australian Geology and Geophysics 7, 1, 63-7.

Morrison, J.L. (1970) A link between cartographic theory and mapping practice : The nearest neighbour statistic. The Geographical Review 60, 4, 494-510.

Morrison, J.L. (1971) Method-produced error in isarithmic mapping. ACSM Cartography Division, Technical Monograph No. CA-5.

Morrison. J L. (1974a) Observed statistical trends in various interpolation algorithms useful for first stage interpolation. The Canadian Cartographer 11, 2, 142-59.

Morrison, J.L. (1974b) The accuracy of isometric maps. U.S. Army Research Office, final report - DA-ARO-D-31-124-71-G45.

Nasca, S.U. (1979) Accuracy volume computations from contour lines. The Australian Surveyor 30, 1, 315-24.

Newton, R. (1968) Deriving contour maps from geological data. Canadian Journal of Earth Sciences 5, 165-6.

Norcliffe, G.B. (1969) On the use and limitations of trend surface models. Canadian Geographer 13, 4, 338-48.

NERC (1974) Surface approximations and contour mapping - User's manual. ECU, NERC, London.

NUMAC (1977) The Ghost graphical output system. NUMAC Computing Service, Newcastle.

Oberlander, T.M. (1968) A critical appraisal of the inclined contour technique of surface representation. Annals, Association of American Geographers 58, 802-13.

Ojakangas, D.R. and Basham, W.L. (1972) Simplified computer contouring of exploration data. Stanford University Publications Geological Science 9, 2, 757-70.

Olea, R.E. (1972) Application of regionalised variable theory to automatic contouring. MSc. thesis, University of Kansas.

Olea, R. (1974) Optimal contour mapping using Universal Kriging. Journal of Geophysical Research 79, 5, 695-702.

Olender, H.A. (1980) Analysis of a triangulated irregular network (TIN) terrain model for military applications. SRI International California, USA.

Oosterom, A. van (1978) Triangulating the human torso. The Computer Journal 21, 3, 253-8.

Ottoson, L. (1978) Establishment of a high density terrain elevation data base in Sweden. 9th International Conference on Cartography, ICA,Maryland.

Pelto, C.R. Elkins, T.A. and Boyd, H.A. (1968) Automatic contouring of irregularly spaced data. Geophysics 33, 3, 424-30.

Petrie, G. (1970) Photogrammetric digitising : Input for data processing. Inter-congress symposium of Commission IV, ISP. Delft.

Petrie,G. and Adam, M.O. (1980) The design and development of a software-based photogrammetric digitising system. Photogrammetric Record 10, 55, 39-61.

Peucker, T.K. (1972) Computer cartography. Association of American Geographers, Resource Paper no.17, Washington.

Peucker, T.K. (1980) The use of computer graphics for displaying data in three dimensions. Cartographica 17, 2, 59-72.

Phoneko, K.M. (1971) Digital terrain model - A photogrammetric approach to selection of route alternatives on basis of earthwork quantities. MSc. thesis, University of Glasgow.

Quenouille, M.H. (1949) Problems in plane sampling. Annals of Mathematical Statistics 20, 355-75.

Rayner, J.N. (1972) The application of harmonic and spectral analysis to the study of terrain. In Chorley, R.J. (Ed.), Spatial analysis in geomorphology. Methuen, London, 283-302.

Reilly, W.I. (1981) Computer-assisted contour mapping. New Zealand Cartographic Journal 11, 1, 5-6.

Renner, W.D., Bahr, G.K. and Compaan, P.C. (1980) A photogrammetric contouring method for radiation therapy. Photogrammetric Engineering and Remote Sensing 46, 3, 321-7.

Rhind, D.W. (1971) Automated contouring - An empirical evaluation of some differing techniques. The Cartographic Journal 8, 2, 145-58.

Rhind, D.W. (1972) One-sided constraints in automated contouring of drift thickness. Geological Society of America Bulletin 83, 2525-32.

Rhind, D.W. (1975) A skeletal overview of spatial interpolation techniques. Computer Applications 2, 3-4, 293-309.

Richardson, J.B. (1975) The automation of fishing charts at the White Fish Authority. The Cartographic Journal 12, 2, 123-7.

Richardus, P. (1976) The precision of areas and volumes derived from medium- and large-scale topographical maps. Photogrammetria 32, 1-13.

Ripley, B.D. (1981) Spatial statistics. John Wiley and Sons, New York.

Robertson, J.C. (1967) The SYMAP programme for computer mapping. The Cartographic Journal 4, 2, 108-13.

Robinson, A.H. (1962) Mapping the correspondence of isarithmic maps. Annals of the Association of American Geographers 52, 4, 414-25.

Robinson, A.H. (1971) The genealogy of the isopleth. The Cartographic Journal 8, 1, 49-53.

Robinson, A.H. and Thrower, N.J.W. (1957) A new method of terrain presentation. Geographical Review 47, 507-20.

Robinson, G. (1972) Trials on trends through clusters of cirques. Area 4, 104-13.

Rogers, A. and Dawson, J.A. (1979) Which digitiser? Area 11, 1, 69-72.

Royal College of Art (1972) Automatic contouring. Paper NERC/RCA (72)1C.

Royal Society (1974) General notes on the preparation of scientific papers. Royal Society, London.

Royle, A.G. , Clausen, F.L. and Frederiksen, P. (1981) Practical universal kriging and automatic contouring. Geo-processing 1, 4, 377-94.

Rozema, W.J. (1969) The use of spectral analysis in describing lunar surface roughness. U.S. Geological Survey Professional Paper no.650-D, D180-8.

Sabin, M.A. (1980) Contouring - A review of methods for scattered data. In Mathematical methods in computer graphics and design. Academic Press, London.

Salomon, K.B. (1978) An efficient point-in-polygon algorithm. Computers and Geosciences 4, 173-8.

Sampson, R.J. (1978) Surface II graphics system. Kansas Geological Survey, Lawrence, Kansas.

Schilcher, M. (1977) A comparison of the accuracy of several contour plots of the Sohnstetten Test Field. Institut für Photogrammetrie der Universitat Stuttgart, Schriftenreihe, 29-50.

Schmidt, W.E. (1969) The Automap system. Surveying and Mapping 29, 1, 101-6.

Schoenberg, I.J. (1946) Contributions to the problem of approximation of equidistant data by analytic functions Quarterly Applied Maths. A, 45-99 and B, 112-41.

Schumaker, L.L. (1969) Approximation by splines. In Greville, T.N.E. Theory and applications of spline functions. Academic Press, New York, 65-85.

Schut, G.H. (1974) Evaluation of some interpolation methods. Proceedings of the Symposium of Commission III. I.S.P. Stuttgart.

Schut, G.H. (1976) Review of interpolation methods for digital terrain models. XIIIth Congress of the International Society for Photogrammetry, Helsinki.

Shaw, E.M. and Lynn, P.P. (1972) Areal rainfall evaluation using two surface fitting techniques. Bulletin of the International Association of Hydrological Sciences 17, 4, 419-33.

Shepard, D. (1968) A two-dimensional interpolation function for irregularly-spaced data. Proceedings of the ACM National Conference 517-24.

Shepard, D.S. (1970) Symap interpolation characteristics. Laboratory for computer graphics and spatial analysis, Harvard Contract No. 68A-2405D.

Shoup, R.G. (1979) Superpaint....the digital animator. Datamation 25, 5, 150-6.

Sibson, R. (1978) Locally equiangular triangulations. The Computer Journal 21, 3, 243-5.

Slootweg, A.P. (1978) Computer contouring with a digital filter. Marine Geophysical Researches 3, 401-5.

Spanner, M.A. (1983) The use of digital elevation model topographic data for soil erosion modelling within a geographic information system. ASP 49th meeting 314-23.

Speakman, J.D. (1980) NOISEMPA: the USAF's computer program for predicting noise exposure around an airport. Aerospace Medical Research Labs, Wright Patterson AFB, Ohio AFAMRL-TR-80-128.

Spicer, J.S. (1979) Property Information Systems for Government. Departments of the Environment and Transport Research Report 30, London.

Sprunt, B.F. (1975) Relief representation in automated cartography: an algorithmic approach. In Davis, J.C. and McCullagh, M.J. (Eds.), Display and analysis of spatial data. Wiley and Sons, London, 113-86.

Stearns, F. (1968) A method for estimating the quantitative reliability of isoline maps. Annals of the Association of American Geographers 58, 2, 590-600.

Stein, A. (1959) Terrain mask angles. Presented at the Geological Society of America Symposium Chicago, Illinois.

Strain, M.B. (1972) Automated contouring from photographic profile data. XIIth Congress of the International Society for Photogrammetry Ottawa.

Surveys and Mapping Branch (1977) A guide to the accuracy of maps Minister of Supply and Services, Canada, cat no. M52-46/1976.

Swain, C.J. (1976) A Fortran IV program for interpolating irregularly spaced data using the difference equations for minimum curvature. Computers and Geosciences 1, 4, 231-40.

Swain, C.J. (1978) Erratum - A Fortran IV program for interpolating irregularly spaced data. Computers and Geosciences 4, 209.

Swindel, G. and van Andel, T.H. (1969) Computer contouring of deep sea bathymetric data. Marine Geology 7, 347-55.

Switzer, P., Mohr, C.M. and Heitman, R.E. (1964) Statistical analyses of ocean terrain and contour plotting procedures. Project Trident Technical Report, Dept. of the Navy Bureau of Ships.

Symbols Committee of the Royal Society (1975) Quantities, units and symbols. Royal Society, London.

Tanaka, K. (1950) The relief method of representing topography on maps. Geographical Review 40, 444-56.

Taylor, P.J. (1977) Quantitative methods in geography. Houghton Mifflin Co., London.

Tempfli, K. and Makarovic, B. (1979) Transfer functions of interpolation methods. Geo-processing 1, 1-26.

Thomas, T.R. (1975) Recent advances in the measurement and analysis of surface microgeometry. Wear 33, 205-33.

Thompson, M.M. and Rosenfield, G.H. (1971) On map accuracy specifications. Surveying and Mapping 31, 1, 57-64.

Thorpe, L.W. (1971) Automatic contouring using Oceanographic data. In International Cartography Association. Automation in Cartography. ICA, Paris, 361-71.

Tobler, W.R. (1965) Automation in the preparation of thematic maps. Cartographic Journal 2, 1, 32-8.

Tobler, W.R. (1969) Geographical filters and their inverses. Geographical Analysis 1, 234-53.

Tobler, W. (1970) Selected computer programs. Department of Geography, University of Michigan, Ann Arbor.

Tobler, W. (1977) Correction to C.J. Swain's program for interpolating irregularly spaced data. Computers and Geosciences 3, 1, 181.

Törlegard, K. (1972) Digital terrain models - General survey and Swedish experiences" Bildmessung und Luftbildwese. 40, 1, 21-30.

Turner, A.K. and Miles, R.D. (1968) Terrain analysis by computer. Indiana Academy of Science Proceedings 77, 256-70.

University of Salford Computing Laboratory (1979) Ginosurf mark 1 user manual. University of Salford Computing Laboratory, Salford.

Unwin, D.J. (1974) How to assess the unknowable. Area 6, 4, 277-8.

Unwin, D. (1981) Introductory spatial analysis. Methuen, London.

Walden, A.R. (1972) Quantitative comparison of automatic contouring algorithms. MSc. thesis, University of Kansas.

Walters, R.F. (1969) Contouring by machine: A user's guide. The American Association of Petroleum Geologists Bulletin 53, 11, 2324-40.

Watson, D.F. (1982) ACORD: automatic contouring of raw data. Computers and Geosciences 8, 1, 97-101.

Watson, G.S. (1957) Analysis of dispersion on a sphere. Royal Astronomical Society Monthly Notices Geophysics Supplement 7, 153-9.

Whitehouse, D.J. and Archard, J.F. (1970) The properties of random surfaces of significance in their contact. Proceedings Royal Society Series A, 316, 97-121.

Whitten, E.H.T. (1970) Orthogonal polynomial trend surfaces for irregularly spaced data. Journal of the International Association for Mathematical Geology 2, 2, 141-152.

Whitten, E.H.T. (1975) The practical use of trend-surface analyses in geological sciences. In Davis, J.C. and McCullagh, M.J. Display and Analysis of Spatial data. Wiley and Sons, London, 282-97.

Wiechel, H. (1878) Theorie und Darstellung der Beleuchtung von nicht gesetzmaessig gebildeten Flaechen mit Ruecksicht auf die Bergzeichnung. Civilingenieur 24, 335-64.

Worster, D.M. (1979) Three-dimensional graph-plotting routines. Computing Service, University of Glasgow.

Worth, C. (1978) Determining a vertical scale for graphical representations of three-dimensional surfaces. The Cartographic Journal 15, 2, 86-92.

Yaglom, A.M. (1965) An introduction to the theory of stationary random functions. Prentice-Hall Inc., Englewood Cliffs, NJ.

Yoeli, P. (1965) Analytical hill shading. Surveying and Mapping 25, 4, 573-9.

Yoeli, P. (1966) Analytical hill shading and density. Surveying and Mapping 26, 2, 253-9.

Yoeli, P. (1975) Compilation of data for computer-assisted relief cartography. In Davis, J.C. and McCullagh, M.J., Display and Analysis of Spatial data. Wiley and Sons, London, 352-67.

Yoeli, P. (1977) Computer executed interpolation of contours into arrays of randomly distributed height-points. The Cartographic Journal 14, 2, 103-8.

Yoeli, P. (1983) Shadowed contours with computer and plotter. The American Cartographer 10, 2, 101-110.

Young, M. (1978) Terrain analysis Program documentation. Dept. of Geography Univ. of Durham. Grant DA-ER0-591-73-G0040, fifth progress report.

Zarzycki, J.M. (1976) Experience with the automated contouring by Gestalt Photomapper GPM-2 for production of 1:50,000 maps. XIII Congress of the International Society for Photogrammetry Helsinki.

Zarzycki, J.M., Harris, L.J. and Linders, J.G. (1975) Topographic cartographic data base. Proceedings of the Commonwealth Survey Officers Conference Cambridge, Paper J2.

APPENDIX 1

COMPUTER INSTALLATIONS AND SOFTWARE UTILISED

A1 COMPUTER INSTALLATIONS AND SOFTWARE UTILISED

A1.1 INTRODUCTION

Throughout this thesis, a heavy reliance has been made on computing resources.

This appendix is concerned with briefly describing the hardware used and exploring the software involved in more depth. The computing facilities will be examined in three broad categories: computer installations, user-developed software and software derived external to this research, whether available commercially or publicly within universities.

A1.2 COMPUTER INSTALLATIONS

One of the aspects of modern-day computing is that users cannot afford to be restricted to one system. Due to the nature of resources and efficiency, information is often best collected on one system (perhaps off-line), may undergo several changes of storage medium (cassette, tape, disk, diskette) and be manipulated by several computers, before the final result is obtained in the final form. The user must be prepared to adopt and adapt new systems and languages to match his immediate and changing needs.

It is therefore not unduly surprising that in the course of this research, one desk-top computer, one micro-computer, one

mini-computer and five main-frame computers have been used.

a. Initially, data were observed on the photogrammetric system in the Department of Geography at the University of Glasgow. This involved a limited amount of programming in BASIC and a working knowledge of the Wang 2200 and its Wang operating system.

b. Additionally, some information was transferred to NUMAC using two systems. The PDP11/34 mini-computer, operated by the Department of Computing at the University of Durham, was used to input and edit programs and data, using both floppy disk input and interactive screen editing. The PDP uses the UNIX operating system.

c. The North Star Horizon micro-computer, in the Department of Surveying at the University of Newcastle, was used to input some isarithm information from a tablet digitiser. The North Star has a modified CP/M operating system, with FORTRAN and BASIC compilers.

d. Most of the computing was performed on the NUMAC

(Northumbrian Universities Multiple Access Computers) IBM 370/168 using the MTS (Michigan Terminal System) operating system. In addition to a wide selection of stand-alone packages, software was written in FORTRAN, ALGOL-W and SIMULA which frequently used the NAG, GHOST, GINO-F and many other subroutine libraries. Local system constraints - even on this sophisticated system - were sometimes of direct relevance to the reasearch. For instance, within MTS at NUMAC, 'in-core' matrices within FORTRAN are limited to 1Mbyte of storage, thus, for example, restricting the use of the program 'MULTI' to data sets smaller than the square root of 125000 if 'DOUBLE PRECISION' (REAL*8) were used (see A1.4.14).

e. The SERC Rutherford IBM 360/195 computers using the OS/360 MVT operating system and the ELECTRIC file handling system were utilised solely to access the SACM package. Data were imported on standard labelled, 9 track IBM tapes and the jobs were run in batch from NUMAC.

f. The University of St. Andrews twin-processor DEC (Digital Equipment Co.) VAX 11/780 using the DEC VAX/VMS operating system was used solely to access the SURFACE II GRAPHICS package. Data were imported on standard labelled, 9 track, IBM tapes and transferred to disk. After processing, plots were obtained locally, and matrices, initially stored on disk, were later transferred to

unlabelled tape for exporting to NUMAC.

g. The University of Aberdeen Honeywell 66/80 using the TSS operating system was utilised to perform early tests on SURFACE II GRAPHICS. It was accessed from the remote site at the University of St. Andrews.

A1.3 EXTERNALLY-DERIVED SOFTWARE UTILISED

During the course of the computing work of this research, several programs, packages, and subroutine libraries were accessed which were not developed by the author. It is important that this software be discussed since the content of the interpolation programs and their chronological development are of significance in any comparison of isarithmic software. Additionally, the software often required additional Input/Output (I/O) software to be developed.

While the majority of the discussion concerns isarithmic or interpolation software, nevertheless some other general libraries (NAG), digitising programs (DIG1), surface roughness programs (GEOD and VECTOR) and general applications programs (SAVDRA) will be considered.

A1.3.1 CONSYS

The CONtouring SYStem ('CONSYS') is a collection of

subroutines for producing isarithm maps only from a regular or irregular grid digital model. The library, written in FORTRAN, was developed by Cederquist (1976) at the Cooley Electronics Laboratory, University of Michigan, and it is the version 1.2, dated 1976, that was accessed at NUMAC.

One of the advantages of CONSYS is that the routines are graphics package-independent. The user must first set the graphics environment within his program and then call the CONSYS initialisation routine, supplying as parameters two routines to move with pen-up or draw with pen-down. CONSYS subsequently uses these move and draw routines to generate the isarithm map. The subroutine library was thus used in conjunction with GHOSYS (see A1.4.9), the calling program designed by the author to draw the isarithm map interactively.

CONSYS was designed to be a simple, yet efficient grid-to-isarithm package and in that respect it succeeds. Isarithms are angular and produced in scans along 'x' from the base to the top of the picture area. The user is responsible for defining the coordinate space and map area, and thus scale is easily controlled. Additionally, a primitive labelling facility exists which the user may invoke by setting a logical variable and supplying the routine to plot a value in a certain location. The author of CONSYS appears to consider the CPU time necessary to produce high quality labelling an extravagance.

A1.3.2 DIG1

The stand-alone program DIG1 (DIGitising 1) was used during

the digitising operations on the Santoni Stereosimplex IIc/Wang 2200 digital photogrammetric system. The program, written in BASIC, was developed and implemented by Adam in 1979 in the Geography Department of the University of Glasgow, where it was used (Petrie and Adam, 1980). The program was interactive and enabled the photogrammetric operator to follow a series of question/answer instructions to select digitising mode, point numbering system and output device. The digitised data were subsequently processed accordingly as they were digitised. In the event, minor changes had to be made to the program to enable tape marks to be written in a compatible form for the post-digitising cassette-to-magnetic tape transfer process.

A 1.3.3 GEOD

The terrain analysis suite of programs 'GEOD' was developed in 1978 at the University of Durham by Young and Young to the specifications of Evans (Young, 1978). The suite of programs, written in FORTRAN and calling SPSS (Statistical Package for the Social Sciences), and MIDAS (Michigan Interactive Data Analysis System), is designed to run under the MTS system at NUMAC where it was accessed. The theory of the program is discussed elsewhere (4.3.3). The input has been kept simple, in the form of a regular grid digital model, which is processed to produce surface characteristics which are displayed by utilising histograms (MIDAS); scattergrams (SPSS); line printer density maps; a plotter produced hachure

map and numerical statistics.

The suite is contained in the source file 'GEOD' which contains all the necessary run commands and parameters and may be edited according to the data set input. Thus the name of the data file, its format, title, grid resolution and size are edited prior to sourcing. The results are of standard form and in general satisfactory although, if necessary, some of the standard parameters may be reselected by consulting the documentation and the source file.

A 1.3.4 GHOST

'GHOST', (GrapHical Output SysTem), is a general purpose graphical subroutine library coded in FORTRAN, and produced by the UK Atomic Energy Authority, Culham Laboratory. The 1978 version was accessed at NUMAC where it is mounted in conjunction with the MTS 'PLOTSYS' library. GHOST, unlike GINO-F, is not device-independent and thus plot descriptor files were NUMAC-specific. It includes a wide selection of graphics drawing routines including grid-to-isarithm routines; a fishnet isometric plotting routine and other graph and three-dimensional drawing routines. To supplement the extant package at NUMAC, a random-to-grid routine mounted in conjunction with the GHOST package at many other universities, was mounted by the author at NUMAC. Random-to-grid interpolation was however kept separate from the isarithm subroutines and the program 'GHOT' (A1.4.10), so-named by the author to distinguish it from the package GHOST, was developed

to perform this aspect.

In general, as a result of flexible routines and superior 'in-house' documentation, GHOST was utilised for all the general graphics throughout this research, with the exception of other specific isarithmic packages. For the purposes of generating isarithms, the program 'GHOSYS' was developed (A1.4.9) to provide a flexible interactive means of producing multi-colour overlain isarithmic plots quickly and efficiently. The 'ISOPLT' routine was used to produce isometric block-diagrams, being easier to use, but less flexible, than PERSYS.

The fact that the GHOST package was used for so much graphics throughout this research suggests its adequacy. Its main flexibility is inherent in the nature of the scaling and windowing, utilising the 'PSPACE', 'CSPACE' and 'MAP' subroutines. Graph drawing (both histograms and point and line graphs) are straightforward, although problems do occur in labelling below the x-axis. The GHOST isarithm routines were found to be both quick and efficient, although this was balanced by somewhat poor labelling procedures and the necessity, as with any FORTRAN program, to pre-define the matrix size.

A1.3.5 GINOSURF And GINO-F

'GINO-F' (Graphical INput/Output - Fortran version) is a general purpose graphical subroutine library, mostly written

in FORTRAN. It was developed at the Computer Aided Design Centre (CADC), Cambridge in 1974, although the version accessed at NUMAC is 2.5B, dated 1979. The interesting feature of GINO-F is that it is not only almost machine-independent, but is also device-independent. While GINO-F is a very flexible graphics package-facilities include two-dimensional and three-dimensional interactive drawing, transformation, selection of symbols, windowing, etc., the package was used purely in its support capacity to 'GINOSURF'.

GINOSURF is a surface display subroutine library, written in FORTRAN, which utilises the GINO-F graphical subroutine library. It was produced in the University of Salford Computing Laboratory in 1979 with the original Mark 1.0 version being mounted at NUMAC. The program performs three basic functions, perspective drawing of regular grids; grid-to-isarithm interpolation and random-to-grid pointwise interpolation (see 2.5.2), although only the last two functions were utilised. As the packages are subroutine libraries, two programs had to be developed to perform random-to-grid interpolation (GINO.RTG) and isarithm interpolation (GINO.CON).

GINOSURF was soon found to be a particularly awkward package to work with. In general, the visual graphical display is adequate, since the border and labelling routines are augmented by the other GINO-F drawing routines to create a powerful system. However, as soon as plots were required to be drawn precisely to scale, for example for overlain isarithm maps, the package was considered inadequate due to a lack of

precise control over scaling. For the purposes of isarithm comparison, the program 'SAVDRA' (A1.3.11) had to be employed to convert the isarithm plot to numerical form and a further program 'RESCAL' used to rescale, list and plot the map out at the required scale and dimensions. While this may be acceptable in a research environment, it is highly undesirable in a production environment.

A 1.3.6 GPCP

'GPCP' (General Purpose Contouring Program) is primarily a contouring program which was developed and marketed by Calcomp in 1971. The source program is in FORTRAN, requiring 32K words of memory, and plotting is performed using Calcomp plotting software. The package was accessed at NUMAC where the original and 1974 supplement are mounted. The original version includes pointwise random-to-grid interpolation; contouring oblique and stereomap production. In 1974, this was supplemented by the inclusion of grid-to-grid operations; global polynomial random-to-grid interpolation; area and volume computation; profiling and additional data posting routines. While anaglyph stereo isarithms are generated elsewhere in this research, GPCP was used primarily to generate isarithms from regular grids (interpolated or observed) and to perform random-to-grid interpolation by utilising both the global polynomial function and pointwise weighted averaging methods (2.5.1).

The package is based on a command structure incorporating precise formatting of data. An interactive program, 'GPCP.IN'

(A1.4.11), was therefore developed to provide the essential data-command-package interface. GPCP.IN was developed to allow for most GPCP options and has therefore subsequently proved invaluable both to the author and as a general service package for other NUMAC GPCP users. It provides a flexible interactive solution to a tricky input situation.

In general, the graphics are flexible, comprehensive and adequate. Scaling is simple and easy to apply, although with particularly large numbers (greater than 99999) some form of user-applied scaling may be required to be performed on the data. Random-to-grid interpolation is limited to two functions only. Conversely, grid-to-isarithm interpolation is flexible and the form of the actual line is widely selectable. Unfortunately, with some of the commands, the manual is not explicit and problems could only be resolved after consulting the source code.

A1.3.7 MINCURV

'MINCURV' (MINimum CURVature) is a random-to-grid interpolation program (2.5.8) written in FORTRAN by Swain (1976) and modified by the author to run at NUMAC with the standard data formats. The program is based on the interpolation method of Briggs (1974) and also includes modifications following Tobler (1977) and Swain (1978). MINCURV consists of a short main routine to input the data and five subroutines to sort the data and perform the interpolation.

A1.3.8 NAG

The 'NAG' (Numerical Algorithms Group) library is an extensive library of subroutines for the solution of a wide variety of numerical analysis problems, for example complex arithmetic, differential equations, equation solution, interpolation and curve and surface fitting and matrix operations. The version accessed at NUMAC was Mark 8, dated 1981, where it was used mainly for the solution of simultaneous linear equations. Throughout the library, which is written in FORTRAN, DOUBLE PRECISION (REAL*8) is employed which necessitates that all data be transformed from REAL*4 before processing. However, this disadvantage is balanced by the ease of use of the library as a sound and thorough documentation supports the main criteria for the selection of the routines, namely their usefulness, robustness, numerical stability, accuracy and speed.

A1.3.9 PERSYS

The 'PERSYS' (PERspective SYStem) collection of subroutines produces 'fishnet' isometric block diagrams from regular and irregular grids. The program version 2.0, dated 1975, was written in FORTRAN by Cederquist at the Cooley Electronics Laboratory, University of Michigan and was accessed at NUMAC. The collection of subroutines is accessible exactly under the same conditions as CONSYS.

PERSYS was only used in conjunction with the surface

characteristics program 'AUTOP' (A1.4.2) to generate isometric correlograms of two-dimensional autocorrelation functions. It was used as it contained the only isometric subroutine available at NUMAC which allowed the user to specify the exact scale and obtain picture co-ordinates from the subroutines. This was necessary as AUTOP required that all plots be similarly sized for future visual comparison. In addition, certain lines were to be overlain on the isometric using the 'GTSCAL' facility. In the event, this was not as flexible as the documentation suggested and PERSYS had to be modified accordingly. Excluding the ability to overlay lines on the isometric plot, PERSYS includes few variations, notably only the ability to produce solid or transparent bases.

A1.3.10 SACM

The 'Surface Approximations and Contour Mapping' (SACM) program suite was developed in 1967 by Assiter at Applications Consultants Incorporated, Houston. The version accessed at the SERC Rutherford Laboratory was mounted in 1974. The full package contains options for random-to-grid interpolation (with or without faulting constraints) using global orthogonal polynomials or pointwise methods; grid-to-isarithm interpolation (graphical or printer output); profiling and perspective plotting. However, only the isarithm and pointwise random-to-grid interpolation were used within this research.

The package is accessed using a command structure involving I/O data in user specified format and filespace. This therefore leads to ease-of-operation although the

documentation is not as adequate as that of SURFACE II GRAPHICS. Scaling of plots and general graphical design is satisfactory - SACM is the only package which selects different pen colours for index isarithms and labels maxima and minima values over the surface. There are however two undesirable features inherent in the random-to-grid interpolation: the package does not consider points lying outside the range of the co-ordinates of the grid corners, thus some edges are always extrapolated and can never be interpolated. Additionally, on the right-hand and top edges, the package generates additional unrequested lines of grid nodes.

A1.3.11 SAVDRA

The name 'SAVDRA' is derived from "SAVDRA" - a 'device' used to produce device independent picture code or pseudo-code within the package GINO-F. In order to produce device-independent transportable plots at NUMAC, the stand-alone program SAVDRA was developed by Hall of the Computing Laboratory, University of Durham where it is mounted. The program is written in SIMULA and was modified to generate a digital plot file of desired format from a NUMAC plot descriptor file. The digital plotfile was subsequently used as input to CONCOR and CONLIN (A1.4.4/5) and was also used when rescaling GINOSURF maps.

A1.3.12 SURFACE II GRAPHICS

The SURFACE II GRAPHICS system, developed at the Kansas Geological Survey in Lawrence, Kansas, is a "computer software system for creation of displays of spatially distributed data" (Sampson, 1978, 1). It was conceived by Sampson and programmed in FORTRAN in 1975, although the 1981 version was accessed at the Computing Laboratory at the University of St. Andrews, where it is mounted in conjunction with GINO-F graphics software. The basic form of graphic display is the isarithm map, although isometric block diagrams (in mono and stereo for pocket, mirror and anaglyph stereoscopic viewing) may also be produced from regular grids only. The program contains an exhaustive selection of random-to-grid interpolation routines (2.5.3) and is unique in providing associated error histograms and other statistical information. Most of the routines were accessed with the exception of the global polynomial interpolation, unavailable in the St. Andrews implementation.

The package is based on a command structure similar to SACM. Scaling and general graphical design could be easily applied within the command structure and the variety of outputs available was ideal. In addition, the documentation is detailed yet concise and explicit. SURFACE II GRAPHICS was thus perhaps the most flexible and complete package considered.

A1.3.13 SYMAP

The SYnographic MAPping program (SYMAP) was produced at the Harvard University Laboratory for Computer Graphics and Spatial Analysis. The source program is written in FORTRAN and version 5.15 (written in 1970 and with some NUMAC modifications) was accessed at NUMAC. The program, designed for line-printer output, has survived primarily as a result of its widespread existence, general flexibility and simple yet efficient documentation and usage. SYMAP may be used to produce three forms of map depicting spatially disposed information, namely, isarithmic conformant (choropleth) and proximal maps. However, due to the nature of the program's output only the pointwise random-to-grid interpolation section within the isarithm routine was used (2.5.5).

Like GPCP, SYMAP is a single input, command structured package, although the structuring is not quite as rigid. It thus suffers from the inherent problem that the data must be pre-formatted and command 'cards' inserted. Additionally, the resultant matrix had to be reformatted from the SYMAP format to the more flexible format used throughout the research. Two short programs were therefore developed, 'SYM.IN' to deal with data/command formatting and 'SYM.OUT' to deal with resultant re-formatting.

SYMAP has several minor problems inherent in its implementation. It must be remembered that SYMAP is designed to produce line-printer maps; thus grid-size in the

random-to-grid interpolation stage is controlled by varying the size of character to be printed. Grid size is defined indirectly and thus creates a certain amount of uncertainty as to the exact mapping of co-ordinates on to printer character cells and vice versa.

A1.3.14 VECTOR

'VECTOR' is a stand-alone element of the program suite written in FORTRAN by Hobson (1967) at Northwestern University to determine surface roughness. The program was mounted at NUMAC where several alterations had to be made to the original source code concerning functions that were not NUMAC FORTRAN-compatible and various I/O formats. Thus the program was established to run smoothly with the standard data format used within this research. The theory of the package is discussed elsewhere (4.4.4). Essentially, data consisted of regular grids. The program, after computation, displayed summary statistics concerning the direction cosines; their vector strength and dispersion; the resultant strike and dip of the mean plane and its mean and standard deviation.

A1.4 SELF-WRITTEN SOFTWARE

Throughout this research a substantial amount of work has been concerned with the development of self-written programs both to perform major aspects of work (isarithmic mapping, interpolation, data/surface characteristic evaluation, and isarithm and matrix comparisons), but also to perform perhaps

more trivial aspects concerned with I/O reformatting. To include information on all the formatting programs would be too bulky and time-consuming - they are too numerous. There are, however three formatting programs worthy of mention, namely REFORM, EFTRAN, and GPCP.IN. In addition, the major programs will be considered in more detail, namely, the data/surface characteristic programs (SA, DTSA, AUTOP, AUFN, SCAVAL and NNA); the random-to-grid interpolation programs (MULTI, BRAILE and GHOT); the random-to-isarithm program (DTC); the subroutine library grid-to-isarithm programs (only GHOSYS will be considered as an example) and the analysis programs (MATANN, CONCOR and CONLIN).

From an early stage in the research, it was realised that in order to obtain a greater amount of programming flexibility, two considerations should be met. Programs are more flexible and simple if basic building blocks or subroutines are common to many programs. Thus a subroutine library (LIB) was evolved throughout the work of the research. As a result of this decision and although several languages were used (BASIC, ALGOL, FORTRAN, SIMULA), most of the programming had to be limited to FORTRAN since until recently at NUMAC, most self-written routines were only accessible by programs written in a similar language.

A1.4.1 AUFN

'AUFN' computes and plots the areal autocorrelation function of a point data set, assumed isotropic, as outlined

in 4.4.1. Input to the program consists of the data in standard format the limits of the area under consideration and a constant to illustrate whether a Geary or Moran coefficient is to be computed. The program then calls the library subroutine 'ARAUT' (described by Henley, 1976). This subroutine was modified to allow for the computation of the Moran or Geary function by the addition of 'NGER' to the parameter list, and the replacement of line 38 by:

```
IF (NGER.EQ.0) SZZ(K)=SZZ(K)+Z(I)*Z(J)*2.0
IF (NGER.EQ.1) SZZ(K)=SZZ(K)+(Z(I)-Z(J))**2
```

A plot of the resultant function, or correlogram, which is specified at user-supplied intervals, is then generated using 'PLOT' which calls GHOST plotting routines. The function is also printed in tabular form.

AUFN consists of a main program and two subroutines. There are no limitations on the program with the exception of the declared maximum data set size (750), which may be redeclared.

A1.4.2 AUTOP

'AUTOP' computes a two-dimensional autocorrelation function of a regular grid digital model of irregular dimension and generates an isometric correlogram within a reference framework. Input to the program consists of a title, the regular grid in standard format (using 'RMATIN'), the dimensions of the input grid, the maximum lags of the resultant function and the viewing perspective parameters for the block diagram. The program then calls 'AUTOCF' (based on

the theory of Cox, 1979) to generate the autocorrelation function (Pearsonian product-moment coefficient), and PERSYS (routines 'PPSET', 'PERS' and 'GTSCAL') to plot an isometric correlogram, before annotating the correlogram with further calls to GHOST routines.

Limitations within the program are few and are primarily due to matrix dimensioning. The maximum grid which may be processed is 120x120 although this may be redeclared. Similarly, the maximum number of lags for the function is 25 in both directions, although if this is changed, X and Y must be similarly altered and AUT set sufficiently large. With relation to the correlogram, the parameters are selected by the user and are utilised indirectly when establishing the annotation lines.

A1.4.3 BRAILE

'BRAILE' performs a patchwise polynomial random-to-grid interpolation following the method of Braile (1978) and discussed earlier (2.5.7). The program contains a short main program to over-ride matrix-transference problems, which calls another short subroutine to input all the other parameters and output the results. This subroutine, 'REST', calls the routine 'GRID' and its associated routines as programmed by Braile (1978) which call 'LINREG' as programmed by Kuester and Mize (1973). The former is responsible for sorting and searching the data for the process of the creation of simultaneous equations for each patch which are solved by the latter.

Within the program there are several limitations. The maximum grid size which may be interpolated is set to 120x120, and the maximum data set which may be input is 1000 points. These upper limits are redeclarable within the coding. Additionally, since the grid is interpolated by sub-dividing it into equi-sized sub-grids, the interpolated grid must be equivalent to an integer multiple of sub-grid nodes. This truncation problem was catered for by including four extra parameters (NNX, NNY, NNXX and NNYY) defined as the maximum and minimum nodes in the x and y directions of the interpolated grid, which represent the actual size of grid required. Finally, it was discovered that the program failed when greater than 100 points were found within any one patch. This was overcome by re-dimensioning (within the Braile 'LSQSRF' routine) matrix XX as (175,65), and YHAT as (175) and setting MI = 175. Thus up to 175 points may now be found within any one patch.

A1.4.4 CONCOR

'CONCOR' derives the intersection, union and associated statistics as a preliminary to computing the coefficient of correlation between computer generated isarithms and the relevant reference isarithms. The program is based on the theory of Court as discussed in 5.3.2. Input to the program consists of the dimensions of the area under interest (in plotfile co-ordinates), the number of isarithm levels to be examined, their starting height, the increment between the

levels and the isarithms in plotfile co-ordinates from lowest to highest level. Output consists of the resemblance matrix and correlation coefficient for each level and a global summary. However, it was a result of problems associated with this program in conjunction with processing a complete isarithmic map in one run (see below), that CONLIN was developed to consider only one isarithmic level per run. The statistics generated were subsequently processed to produce the summary diagrams and tables (see Chapter 7).

Having input and initialised the relevant information in the main routine, which also produces the final global summary statistics, control is passed to 'CO' which is the routine responsible for performing the main isarithmic level loop. Within CO, the co-ordinates of the first level for the isarithms under interest and the reference isarithms are input. Five separate operations must now be undertaken.

a. The intersection points of the INTEREST and REFERENCE isarithms are computed using 'GETINT' which utilises 'INTER' (derived from Baxter, 1976, 195).

b. The intersection points must be matched and preliminary cross-referencing labels established. This is performed utilising 'MATCH' and 'MISSED', the latter subsequently calling 'EDGE' to deal with the border of the area.

c. Additional point labels are established using 'INTOK' for all the non-intersection points. This is performed by considering the point-in-polygon results, generated by 'PTPOLY' which in turn utilises 'PREPLY', 'SORT', 'INCLUD', 'ORDER' and 'INOT'. 'PTPOLY' and the associated routine 'PTPOL2' are adaptations of the routines proposed by Salomon (1978).

d. Having established the points and their labels, 'U' and its associated routines 'GUTS' and 'ENDSEG' are called to derive the co-ordinate strings for each of the union areas.

e. The areas of the union strings are computed using 'CONARE' which calls the routine 'AREACM' (after Hall, 1976).

Summary statistics are computed for each level within 'CO' before recommencing the isarithmic level loop.

It should be noted from the outset that, while 'CONCOR' is based on simple logic, nevertheless the program suite involves 22 routines in approximately 1800 lines of source code and is thus very sensitive. As regards programming, only two limitations exist. Within any one level, a maximum of 25 separate isarithms may exist containing a total of 1000 points including intersection points per level. Both these limitations may be redeclared by the user. More important is

the limitation of time. Because the program necessitates the frequent evaluation of a point-in-polygon routine, which is invariably time-consuming (although the méthod used was considered superior to other methods considered), the total run-time for a whole isarithmic map (9 levels) frequently exceeded the maximum allowable limit at NUMAC (1200 cpu seconds). In such a case, the map was evaluated by considering each level separately.

A1.4.5 CONLIN

'CONLIN' compares the relative geometric accuracy and morphological trueness of isarithms. The program, which is based on the theory of Lindig (1956a, 1956b) as discussed in Chapter 5, computes the Position, Elevation, Direction and Curvature errors for individual computer-generated isarithms within a model. Input to the program consists of a title (the isarithm level); the dimensions of the area under interest (in both real and plotfile co-ordinates); the isarithms (in plotfile co-ordinates from 'SAVDRA'); a gradient matrix for the model area and additional parameters R and TOL used in the computation. Output, which consists of the summary statistics, is sent to two units - as a 'quick-look' facility in addition to an abridged version which is stored for each level and subsequently processed for the whole model to produce the summary diagrams (see Chapter 7). Having input and rescaled the relevant data using 'READIN' and 'REGRAD', the program generates the equations of the tangents at every point on the

REFERENCE isarithm and the normals to the tangents on the isarithm under INTEREST (using 'GETXY' and 'GETX', which in turn call 'EQUAT' to derive the unknowns). By solving for each pair of tangent and normal equations (using 'INTER'), the nearest point may be found on the REFERENCE isarithm to each INTEREST point along the normal through each point. The set of such points is then used with the points on the isarithm under INTEREST to compute Position error (distance along the normal), Elevation error (Position error x tangent (gradient) using 'ELEV'), Direction error (difference in orientation between two consecutive points on the isarithm under INTEREST and the respective 'nearest' points on the REFERENCE isarithm) and Curvature error (using 'CURV'). Within CURV a value for R (the reference radius of curvature) is used. This was set to 500.0 throughout all runs, although it was programmed as an input parameter. Finally, 'ASSYM' is called to tidy up the errors produced and remove any gross errors. This first removes any errors at points with a Position error greater than TOL (an input parameter set to 200.0 for INCH and LIAN and 70.0 for FORV). Subsequently, the standard deviation for each distribution is computed (using 'RSDEV') and errors greater than 2.5 standard deviations removed. The final summary statistics are generated.

There are several limitations to the program mainly concerned with matrix sizes, although with care these may be easily re-defined. The gradient matrix has maximum dimension (100.100) and the maximum number of points that may be processed in any isarithm level at one time is set to 1000. Within the program, isarithms may be closed loops or open line

segments. While the initial aspirations for the program were to compute the statistics for all levels in one run, this was subsequently considered too restrictive; it required bulky matrices in some instances and due to the limitations of job priorities with run-time at NUMAC was also considered undesirable. The program was therefore developed for computing the statistics for each isarithm line within each level separately.

A1.4.6 DTC

The program 'DTC' produces an isarithm map from random point data by utilising a Delaunay triangulation. After the random three-dimensional data are input to the program, a Delaunay triangulation is generated, utilising TILE - a series of routines written by Green and Sibson (1978), which may be subsequently used to derive a Dirichlet (Voronoi, or Thiessen) tesselation and/or an isarithmic map using user-desired isarithm levels. The isarithms, produced by direct linear interpolation from the triangulation, may be left angular or may be smoothed. The map, tesselation or triangulation may be user-scaled.

The Delaunay triangulation, generated from calls to TILE, optimises to an equiangular triangulation within a 'window' or bounded polygonal region. The window is thus an important aspect to be defined as the triangulation may vary at the edges depending on its definition. This is the only feature that should prevent a unique triangulation solution given a

fixed input data set.

The output from the triangulation must be processed to form a more flexible system for isarithm generation, using 'FORM'. This is required to define the triangles, their neighbours and their maximum and minimum values for the isarithm interpolation stage. Isarithms are traced using 'CONTUR' and may be smoothed using 'CRVFIT' (Akima, 1974a) once the whole of an isarithm has been generated. Additionally, a routine 'PLTPT' is supplied which may be used to plot and label points. The pre-compiled FORTRAN triangulation/tesselation subroutine library 'TILE' contains the following subroutines;

a. TBOX - defines the constraints of the tesselation,

b. TILE - derives the tesselation,

c. PRINT - produces a tabulation of constraint coefficients, contiguity lists and point coordinates,

d. STATS - produces a tabulation of various tile statistics - area, perimeter etc.,

e. PLTDIR - plots the tesselation triangulation, or datapoints.

FORM, CONTUR, CRVFIT and PLPT are in the research

subroutine library and call GHOST plotting routines.

There are three known limitations. The declaration parameters may be altered to accept a larger number of points. The input data must be sufficiently random: in the event of a regular grid being input, the tesselation generated causes ambiguities as each node considers the four surrounding nodes to be contiguous, thus preventing a triangulation from being established. Finally, a feature of the isarithm routine is that the border must be carefully defined. For the purposes of this research, where isarithms were traced through a precisely defined rectangular area, the border component of the data sample was used to prevent the triangles from overlapping the area of interest.

A1.4.7 DTSA

The program 'DTSA' computes the surface areas of individual triangular cells formed from point data, their resultant summary statistics and plots a histogram of the distribution with the summary statistics. It combines a small surface area computation section sandwiched between the data-to-triangulation section of DTC and the statistical section of SA. Input is as for DTC. After a call to 'FORM', the individual cell surface areas are computed and compared with the cell plane areas using 'DIST3D', 'DIST2D' and 'SURFAR'. The summary statistics are then computed using 'STDDEV'(to compute standard deviation) and 'SKKU' (to compute skewness and kurtosis), and printed. The histogram and summary statistics are then plotted using 'COMHIS' and 'HIST' which

call GHOST graphics library routines.

The only limitations to the program are those resulting from the use of the 'TILE' suite and are thus similar to those for 'DTC'. FORM, DIST3D, DIST2D, SURFAR, STDDEV, SKKU, COMHIS and HIST are all in the subroutine library.

A1.4.8 EFTRAN

Throughout the research, a considerable number of small input/output programs were written to manipulate matrices. 'EFTRAN' is one such program written to convert GPCP matrix output in 'E' format to the standard 'F' format used throughout the research. It is therefore typical of many such programs eg. 'SYM.OUT', and is thus included as an example.

EFTRAN is interactive initially requesting the dimensions of the grid to be reformatted. Once input the program reads in, converts and outputs the grid. There is only one limitation on the program - the maximum size of the grid (120,120), which is changeable by the user.

A1.4.9 GHOSYS

The interactive program 'GHOSYS' performs grid-to-isarithm interpolation using either regular or irregular grids, with the option of selecting the routines available in GHOST or CONSYS. Within CONSYS, the GHOST 'POSITN' and 'JOIN' routines are called within 'CONSET' to provide the move and draw routines for 'CONTUR' to trace the isarithms. Both packages are subroutine libraries; therefore the GHOSYS program reads

in data and produces the isarithm map according to user-specified parameters. Maps may be drawn at user-specified scale with other maps and/or digitised line work overlain. In all cases, input data must be in the author's standard research format.

The isarithms are drawn by calling one of five subroutines,

a. CONTUR - is the only CONSYS routine which produces angular isarithms from a regular or irregular grid;

b. CONTRA - is the GHOST routine which produces smoothed isarithms from a regular grid;

c. CONTIA - produces a similar result from an irregular grid;

d. CONTRL - is the GHOST routine which produces angular isarithms from a regular grid;

e. CONTIL - produces a similar result from an irregular grid.

GHOSYS consists only of a main program. There are no limitations on the program, with the exception of the grid size which is user selectable, but it has a changeable maximum default size of 120x120. Additionally, if a title is desired,

this is plotted vertically in a 1 inch band on the left hand side of the plot.

A1.4.10 GHOT

'GHOT', as outlined in 2.5.6, is a random-to-grid interpolation program which calls the interpolation routines REGRID and DWLSQ1 of Brodlie. These routines were released to University users of GHOST and are often mounted in conjunction with this package. The routines were acquired from the University of Glasgow Computing Service where they are mounted in conjunction with GHOST (Worster, 1979). The main program is brief. After reading the various parameters necessary and the required data using 'RPTIN', it calls the routine 'REGRID' and thereafter prints the results using 'RMATOT'. REGRID generates the size of the search radius from the input data and parameters and calls 'DWLSQ1' to perform the pointwise least squares polynomial interpolation. All routines are to be found in the subroutine library.

Excluding any problems associated with REGRID and DWLSQ1, the only limiting factor is the size of the arrays which are changeable with maximum defaults '(120 120)' for the grid, and '(4000)' for the point data. It should be noted that problems may occur with the search radius in areas of sparse data, where it may not be satisfactory and provide too few points and therefore an ill-conditioned least squares fit.

A1.4.11 GPCP.IN

The interactive program 'GPCP.IN' is a service program

designed specifically to perform the data reformatting which is a pre-requisite to running GPCP. GPCP as discussed earlier is a command-based program which requires that all data and option commands be structured within the same file. Thus, throughout the program, the user must not only select the necessary routines and parameters that are required, but also must specify the data type (grid or point data) and the format in the case of the point data. The various options available, which are generally selected by a YES/NO response to prompts, represent a fairly extensive subset of the total available input to the package and as well as standard options, include the TREND, stereo, boundary and line drawing facilities. The relevant commands are output as the program progresses, the program terminating only when the commands are completed with a 'STOP' card. This command file may now be directly input to GPCP.

GPCP.IN consists of a main program and subroutines for selecting newpens (NEWPEN); line drawing (CONPLO) and boundary insertion (BOUNDY). There are no limitations on the program with the exception of the grid size which is user selectable, but has a changeable maximum default size of 120x120. The CONPLO and BOUNDY routines are set to read data in 'free format'.

A1.4.12 LIB

'LIB' is the subroutine library established and developed throughout the course of the research to facilitate ease of programming, compaction of programming, portability of

subroutines (both by the author and other interested users) and the development of a comprehensive system for any future research. In that respect the library was not envisaged as a collection of routines each of which had to be used in several programs before they could be accepted in the library. Instead, if a routine was potentially important or desirable and it was in 'perfect running order', it was removed to the library. Latterly, the library contained sixty routines.

A1.4.13 MATANN

'MATANN' was the main program used to analyse statistically the grid products generated by the various random-to-grid interpolation methods. Much of the discussion on the choice of statistical techniques to be utilised may be found in Chapters 5/6. Essentially, the program generates standard parameters of error and correlation between each interpolated and observed (reference) grid. The error is correlated additionally against grids of local terrain gradient, distance to the nearest data point, and number of data points within a prescribed distance from the grid node. In all cases, point error above a certain tolerance (approximately 40% of the range of data) is ignored and a total of points 'not estimated' augmented.

The program places a heavy reliance on library subroutines. The various grids are read in using 'RMATIN' and 'IMATIN' along with the title and tolerance. Error grids are then generated and using 'COR2' and 'TRANS' (to transform the distributions) the various grid correlation coefficients are established. All other statistical parameters are established

utilising 'RSDEV' (standard deviation), 'RSER2' (standard error) and 'SKKUR' (skewness and kurtosis). Best fit lines are computed for each of the error/grid comparisons using 'CLASS' and 'EQNLIN' (fits line to classified data). The program consists of the main routine CLASS (sets random distributions into classes) and EQNLIN; the other routines are found in the library.

There are two limitations to the use of the program. The program is designed for a 49x49 grid and 12 classes within the best-fit line section. If an alternative number of grid nodes or classes are required then the declaration parameters must be redefined. Additionally, inherent in any least squares computation is that any matrices must not be ill-conditioned. Thus problems may occur if many classes contain no elements, requiring class limit reselection.

A1.4.14 MULTI

The program 'MULTI' generates a regular grid from random data by the theory of multiquadric analysis (2.5.9). The main program is designed purely to overcome matrix transference problems and, after the declarations, calls a subroutine 'REST'. This in turn reads in all the other parameters and data ('RPTIN'), calls the main routine 'MQA' to perform the multiquadric analysis and prints out the results using 'RMATOT'. The routines RPTIN, RMATOT and MQA are defined in the subroutine library.

The routine MQA, in performing the interpolation, requires the solution of a series of simultaneous equations

representing a series of quadratic cones centred on each data point. This is performed by utilising a NAG routine (F04ARF) and creates two associated problems. Data points must be discrete, but this is generally normal in most random-to-grid interpolations. Alternatively, as each data point has a cone which is fitted to all other data points, a large, DOUBLE PRECISION matrix is required, resulting in serious storage limitations in its use. Thus the matrix declarations in the main program are set to the largest allowable size at NUMAC, approximately 340x340 for the A matrix. The declaration parameters for the grid size set at 120x120 may be redefined.

A1.4.15 NN

'NN' computes various Nearest Neighbour Statistics associated with the data and random-to-grid interpolation. Primarily utilising 'NNA', it computes the Nearest Neighbour Analysis statistic for the data distribution. For the purposes of random-to-grid error statistics, the routine 'NNGRID' is subsequently called to compute the distance to the nearest data point and the number of data points within a fixed tolerance of each interpolated grid node. Data are input and resultant matrices printed using 'RPTIN', 'RMATOT' and 'IMATOT'. All subroutines are included in the library.

The only limitations are the sizes of the matrices, presently set to 340 (maximum number of data points) and (120,120) (maximum grid size), which may be reselected. The program is based on well-established theory (see for example Taylor, 1977, 156-62).

A1.4.16 PTDIST

The program 'PTDIST' analyses the pattern of the data sets considered in this research utilising distance-based measures which assume a Euclidean geometry approach (see 3.4.5.2). After computing all the inter-data point distances, the program computes the mean centres and standard deviations within the main program. In addition, the distances to the first, second- and third-order neighbours for each data point are stored, and the histograms printed (using the library routines 'COMHIS' and 'RHISTOT') and distribution statistics derived (using the library routines 'STDDEV', 'SKKU' and 'RQUART').

This program is based on the relevant discussion within Unwin (1981). The only limitation to its use is concerned with data set size (provisionally set to accept a maximum of 340 points).

A1.4.17 QA

The program 'QA' performs a quadrat analysis of the various point data sets used in this research using the quadrat censusing technique (see 3.4.5.1). The program contains a main routine and calls the library routines 'RPTIN' (to read in the data); 'STDDEV'; 'SKKU' and 'RQUART' (to compute the general distribution statistics) and 'COMHIS' and 'RHISOT' (to compute and print the distribution). The main quadrat analysis specific statistics are computed within the main program.

In evaluating the latter statistics, 'QA' is based on the

theory expressed in Unwin (1981) and more specifically Ripley (1981). For a user-specified grid dimensions, resolution and origin the program computes the mean, variance, ratio, index of clustering, index of cluster frequency, index of mean crowding and index of patchiness. The only limitations of the program are those concerned with the number of data points input (maximum 340) and the maximum number of quadrats considered (4000).

A1.4.18 REFORM

The interactive program 'REFORM' is a service program designed specifically to perform the essential task of deriving sub-grids and nodal point data from regular grids. Once the data grid is input in user-specified format, an interactive discussion takes place during which the user may select the general format of the output grid, whether it is to have a terminator and whether the x, y coordinates are to be output. The relevant data are output and the user may request the same data set or a new data set to be processed, or the program to terminate. REFORM consists solely of a main program. There are no limitations on the program with the exception of the grid size which has a changeable maximum default size of 120x120. While the data input formats are user-definable, the output formats are defined in the program to be the standard format used throughout the research.

A1.4.19 SA

'SA' computes the surface areas of the individual cells of

a regular grid terrain matrix, their resultant summary statistics and plots a histogram of the distribution with the summary statistics. The input to the program consists of the grid (using 'RMATIN'), its dimensions, resolution and title. The area of each grid cell is computed by dividing the square cell into two triangles and computing the area of each using the routine 'SURFAR'. This is compared with the plane area of the cell and the ratio surface area/plane area computed. The summary statistics are then computed using the routines 'TRANS' (to transform the distribution), 'STDDEV' (to compute standard deviation), 'SKKU' (to compute skewness and kurtosis) and 'RQUART' (to compute the quartiles), and printed. A histogram and summary statistics are then plotted using COMHIS and HIST. The routines RMATIN, SURFAR, SKKU, STDDEV, RQUART, TRANS COMHIS and HIST are in the subroutine library and call GHOST graphics routines.

There are only two limitations to the program. The size of the input grid is user-definable, but has a changeable maximum default size of 120x120. The DOUBLE PRECISION matrix 'VOL' contains the cell areas and has a changeable maximum default size of 5000. VOL must be at least of size (NX-1)*(NY-1).

A1.4.20 SCAVAL

'SCAVAL' computes and tabulates the scale - variance components of a regular grid following the theory of Moellering and Tobler (1972), see 4.4.5. The program contains a main routine and subroutine and calls the library routines 'RMATIN' and 'RMATOT' to read in and print out matrices.

Having performed the necessary initialisations (title, dimension, resolution and grid input), the mean and deviation matrices for the first level are computed using 'CALC' and the summary statistics generated. The mean matrix 'M' is then used as the new level and the process repeated until only one global mean is derived. Throughout, the statistics are printed to units 6 or 7.

The iterative process therefore enforces one limitation, that of a square regular grid of size equal to a power of two. Additionally, a default maximum grid size of 120x120 is set, although this is changeable.